DOCUMENT RESUME

ED 449 354	CE 081 270

AUTHOR	Kirsch, Irwin; Yamamoto, Kentaro; Norris, Norma; Rock, Donald; Jungeblut, Ann; O'Reilly, Patricia; Berlin, Martha; Mohadjer, Leyla; Waksberg, Joseph; Goksel, Huseyin; Burke, John; Rieger, Susan; Green, James; Klein, Merle; Campbell, Anne; Jenkins, Lynn; Kolstad, Andrew; Mosenthal, Peter; Baldi, Stephane
TITLE	Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-2001-457
PUB DATE	2001-01-00
NOTE	648p.; Project Officer: Andrew Kolstad.
AVAILABLE FROM	ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827 (Toll Free), Fax: 301-470-1244, TTY/TDD: 800 437-0833, E-mail: edpubs@inet.ed.gov, Web site: http://www.ed.gov/pubs/edpubs.html.
PUB TYPE	Guides - Non-Classroom (055) -- Reports - Research (143) -- Tests/Questionnaires (160)
EDRS PRICE	MF03/PC26 Plus Postage.
DESCRIPTORS	Adult Basic Education; *Adult Literacy; *Data Processing; Educational Research; Illiteracy; Literacy Education; *Material Development; *National Surveys; Prisoners; Research Administration; Research Design; Research Methodology; Research Problems; Sampling; Scaling; State Surveys; Statistical Analysis; Validity; Weighted Scores
IDENTIFIERS	*National Adult Literacy Survey (NCES)

ABSTRACT

Chapter 1 of this report and user's manual describes design and implementation of the 1992 National Adult Literacy Survey (NALS). Chapter 2 reviews stages of sampling for national and state survey components; presents weighted and unweighted response rates for the household component; and describes non-incentive and prison sample designs. Chapter 3 addresses weighting procedures. Chapter 4 describes NALS's conceptual framework and development of the background questionnaire and literacy tasks. Chapters 5 and 6 document field operations for the household and prison surveys. Chapters 7-9 detail data processing; the missing data procedures; and models and procedures used to scale NALS results, estimate respondents' proficiencies, and conduct statistical analyses. Chapter 10 presents a summary of literature on use of monetary incentives in survey research, experimental features of the field test and national study, and research results. Chapters 11-15 describe statistics, components of variance, and statistical methods used to derive outcome estimates; estimate and test discriminant validity of the three literacy scales from the perspective of correlation or covariance; summarize establishment of NALS literacy levels; explore importance of the response probability convention in reporting prose literacy results; and discuss use of SPSS (Statistical Program for the Social Sciences) and SAS (Statistical Analysis System) for analyzing NALS data. Appendixes include 98 references, additional data and notes, and instruments. (YLB)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 449 354

ATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report January 2001

Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

CE 081 270

U.S. Department of Education
Office of Educational Research and Improvement NCES 2001–457

Customer Survey of OERI Publication Users

To help us improve future editions of this publication and give you better customer service, we would appreciate your comments on this survey form. Please check the appropriate box(es) below for each question. Responses will be kept completely confidential. You may return the survey by mail or FAX. It can be folded and taped closed to allow mailing to the address listed on the reverse side of this form, or it can be returned by FAX to 202–219–1321. **Many thanks for your customer feedback—it is very important to us!**

1a. Name of publication

1b. Publication number NCES 2001–457

1c. Author name

2. How did you receive a copy of this publication?

☐ Bought it
☐ Borrowed it
☐ Mailing list membership
☐ Telephone request
☐ Internet request
☐ Other (please describe)

3. Was this publication easy to get?

☐ Very ☐ Somewhat ☐ Not at all

4. How did you find out about this and other OERI publications? (Check all that apply.)

☐ Conferences
☐ Journal articles
☐ Teacher/educator
☐ Professional associations
☐ Internet (WWW)
☐ Publication announcement
☐ Received in mail
☐ OERI staff contact

5. For what purposes did you use this OERI publication? (Check all that apply.)

☐ Planning
☐ Policy or legislation
☐ Administrative decisions
☐ Teaching, class material
☐ Research/analysis
☐ General information
☐ Writing news articles, TV or radio material
☐ Marketing, sales, or promotion
☐ Other (please describe)

6. Did the publication help you accomplish whatever you needed it for?

☐ Yes ☐ No ☐ Partially

7. What is your occupation?

☐ Parent ☐ Teacher ☐ Administrator
☐ Librarian ☐ Researcher ☐ Statistician
☐ Journalist/writer ☐ Policy Analyst ☐ Student
☐ Program Planner
☐ Other (please specify)

8. How could this OERI publication (or other OERI publications) better meet your needs? (Check all that apply.)

☐ More important topics in education
☐ More timely release of data
☐ More text introductions to each section
☐ More research statistics
☐ Shorter reports (less than 10 pages)
☐ Other (please describe)

9. Overall, how satisfied are you with this product?	**Very Satisfied**	**Satisfied**	**Dissatisfied**
	☐	☐	☐
a. Comprehensiveness of information	☐	☐	☐
b. Clarity of writing (readability, interpretability)	☐	☐	☐
c. Clarity of presentation (e.g., tables, charts)	☐	☐	☐
d. Timeliness of information	☐	☐	☐
e. Accuracy of information	☐	☐	☐
f. Clarity of technical notes	☐	☐	☐
g. Usefulness of resources and bibliography	☐	☐	☐
h. Organization	☐	☐	☐
i. Length	☐	☐	☐
j. Format	☐	☐	☐

PAPERWORK BURDEN STATEMENT
Office of Education Research and Improvement (OERI)
Publication Customer Survey

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 1880-0529. The time required to complete this information collection is estimated to average 10 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have any comments concerning the accuracy of the time estimate(s), suggestions for improving this form, or concerns regarding the status of your individual submission of this form, write directly to: **P. Quinn, Room 204, Media and Information Services, OERI, U.S. Department of Education, 555 New Jersey Avenue NW, Washington, DC 20208-5570.**

BEST COPY AVAILABLE

OERI Publication Customer Survey
Media and Information Services
U.S. Department of Education
Washington, DC 20202

Official Business
Penalty for Private Use, $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 012935 WASHINGTON DC

POSTAGE WILL BE PAID BY U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education
Mail Code: 5570
400 Maryland Avenue, S.W.
Washington, D.C. 20277—2935

Fold on line—TAPE CLOSED—DO NOT STAPLE

10. Do you have any suggestions regarding the content or format of future editions of this publication or other comments?

MIS 1999-6532

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report January 2001

Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey

Irwin Kirsch, Kentaro Yamamoto, Norma Norris, Donald Rock, Ann Jungeblut, and Patricia O'Reilly—*Educational Testing Service*

Martha Berlin, Leyla Mohadjer, Joseph Waksberg, Huseyin Goksel, John Burke, Susan Rieger, James Green, and Merle Klein—*Westat, Inc.*

Anne Campbell—*Diné College*

Lynn Jenkins—*Wordsworth Writing and Editing*

Andrew Kolstad—*National Center for Education Statistics*

Peter Mosenthal—*Syracuse University*

Stéphane Baldi—*American Institutes for Research*

Andrew Kolstad
Project Officer
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement **NCES 2001–457**

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
1900 K Street, NW
Washington, DC 20006–5651

January 2001

The NCES World Wide Web Home Page is: ***http://nces.ed.gov***
The NCES World Wide Web Electronic Catalog is: ***http://nces.ed.gov/pubsearch/index.asp***

Suggested Citation
U.S. Department of Education. National Center for Education Statistics. *Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey*, NCES 2001–457, by Irwin Kirsch, Kentaro Yamamoto, Norma Norris, Donald Rock, Ann Jungeblut, Patricia O'Reilly, Martha Berlin, Leyla Mohadjer, Joseph Waksberg, Huseyin Goksel, John Burke, Susan Rieger, James Green, Merle Klein, Anne Campbell, Lynn Jenkins, Andrew Kolstad, Peter Mosenthal, and Stéphane Baldi. Andrew Kolstad, Project Officer. Washington, DC: 2000.

For ordering information on this report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

or call toll free 1–877–4ED–Pubs.

Content Contact:
Andrew Kolstad
(202) 502–7374

TABLE OF CONTENTS

CHAPTER 1: THE NATIONAL ADULT LITERACY SURVEY: AN OVERVIEW 1
- 1.1 Introduction 1
- 1.2 Defining Literacy 3
- 1.3 The Sample 3
- 1.4 Weighting 4
- 1.5 The Survey Instrument: Measuring Literacy 5
- 1.6 Field Operations 7
- 1.7 Data Processing and Missing Data 8
- 1.8 Scaling and Proficiency Estimates 9
- 1.9 Establishing Literacy Levels 10

CHAPTER 2: SAMPLE DESIGN 11
- 2.1 Overview 11
- 2.2 Sampling for the National Component 13
  - 2.2.1 First-Stage Sample 14
    - 2.2.1.1 Westat's master sample of PSUs 14
    - 2.2.1.2 Selecting the sample of PSUs for the national component 14
  - 2.2.2 Second-Stage Sample—Selecting Census Blocks (Segments) 21
    - 2.2.2.1 Measures of size and sampling rates 22
    - 2.2.2.2 Minimum segment size 23
    - 2.2.2.3 Segment sample selection 24
    - 2.2.2.4 TIGER maps 24
    - 2.2.2.5 Listing sample segments 25
  - 2.2.3 Third-Stage Sample—Selecting Housing Units 26
    - 2.2.3.1 Within-segment sampling rate 26
    - 2.2.3.2 Overall probabilities of selection 28
    - 2.2.3.3 Procedures for selecting missed structures and missed dwelling units 28
  - 2.2.4 Fourth-Stage Sample—Selecting Persons Age 16 or Older 29
- 2.3 The Non-Incentive Sample 30
- 2.4 Sampling for the State Literacy Surveys 30
  - 2.4.1 Sample of PSUs 31
  - 2.4.2 Sample of Segments 31
  - 2.4.3 Sample of Housing Units 32
  - 2.4.4 Sample of Persons 32
- 2.5 Weighted and Unweighted Response Rates 32
- 2.6 Sampling for the Prison Survey 34
  - 2.6.1 Sample of Correctional Facilities 34
    - 2.6.1.1 Sampling frame and selection of correctional facilities 34
  - 2.6.2 Selection of Inmates Within Facilities 37

CHAPTER 3: WEIGHTING AND POPULATION ESTIMATES 39
- 3.1 Goals of Weighting 39
- 3.2 Calculating Sample Weights for the Household Population 41
  - 3.2.1 Preliminary Steps in Weighting 41
  - 3.2.2 Computing Base Weights 43
  - 3.2.3 Nonresponse Adjustments and Poststratification 44
  - 3.2.4 Compositing Data from the National and State Components 45
    - 3.2.4.1 Composite estimation procedure 45

TABLE OF CONTENTS — CONTINUED

3.2.4.2 Deriving the PSU design effect 48
3.2.4.3 Estimating composite factors 49
3.2.5 Computing Final Weight—Poststratification Through Raking Ration Adjustments 54
3.3 Replicated Weights for Variance Estimation in the Household Population 55
3.3.1 Household Sample Replication for the National Component 57
3.3.2 Household Sample Replication for the State Component 57
3.3.3 Final Household Sample Replication for the National and State Components 58
3.4 Calculating Sample Weights for the Prison Population 58
3.4.1 Computing Inmate Base Weights 58
3.4.2 Nonresponse Adjustments 62
3.4.2.1 Facility nonresponse adjustment 62
3.4.2.2 Inmate nonresponse adjustment 63
3.4.3 Poststratification Procedures 64
3.4.4 Final Inmate Weights 67
3.5 Replication Weights for Variance Estimation in the Prison Population 68

CHAPTER 4: DEVELOPMENT OF THE SURVEY INSTRUMENTS 70
4.1 Conceptual Framework 70
4.2 The Scope of the Background Questionnaire 72
4.2.1 General and Language Background 73
4.2.2 Educational Background and Experiences 73
4.2.3 Political and Social Participation 74
4.2.4 Labor Force Participation 74
4.2.5 Literacy Activities and Collaboration 75
4.2.6 Demographic Information 75
4.2.7 Prison Survey Background Questionnaire 76
4.2.8 Spanish Versions of the Questionnaires 76
4.3 Development of the Simulation Tasks 76
4.3.1 Organizing Framework for Task Development 77
4.3.2 Materials/Structures 77
4.3.3 Adult Contexts/Content 78
4.3.4 Processes/Strategies 79
4.3.5 Task Difficulty 87
4.3.6 Development of Scoring Guides 88
4.3.7 Assembling the Tasks for Administration 89

CHAPTER 5: THE HOUSEHOLD SURVEY 93
5.1 Overview 93
5.2 Listing 94
5.2.1 Staff Organization for Listing 94
5.2.2 Training Listers 95
5.2.3 Listing Materials 95
5.2.4 The Listing Operation 96
5.2.5 Quality Control Procedures 97
5.2.5.1 Quality control of listing sheets 97
5.2.5.2 Quality control of the listing operation 97
5.3 Data Collection Instruments and Interviewer Materials 99
5.3.1 The Screener 99
5.3.2 Interview Guides for Exercise Booklets 101
5.3.3 Non-interview Report Forms 101

8

TABLE OF CONTENTS — CONTINUED

5.3.4 Interviewer Manuals 104
5.3.5 Field Aids 104
5.3.5.1 Aids used for locating and contacting respondents 104
5.3.5.2 Aids used for obtaining respondent cooperation 105
5.3.5.3 Aids used during the interview 106
5.4 Field Organization and Training 106
5.4.1 Field Organization 106
5.4.1.1 Lines of responsibility 106
5.4.1.2 Interviewer recruitment 107
5.4.2 Training 109
5.4.2.1 Supervisor training 109
5.4.2.2 Interviewer training 110
5.4.2.3 Editor training 114
5.5 Field Operations 114
5.5.1 General Approach to the Field Effort 115
5.5.2 Schedule and Production 116
5.5.3 Reporting Systems 117
5.5.3.1 Automated Survey Control System (ASCS) 117
5.5.3.2 Interviewer reports to the supervisor 118
5.5.3.3 Supervisor reports to the home office 118
5.5.3.4 Home office staff reports to ETS and to NCES 119
5.6 Quality Control of Data Collection 119
5.6.1 Introduction 119
5.6.2 Editing 119
5.6.3 Validation 121
5.6.4 Observation 122
5.6.5 Supervisor Observations 123
5.7 Response Rates 123
5.7.1 Reasons for Non-response 125
5.7.2 Characteristics of Non-respondents 126
5.7.3 Discussion 128

CHAPTER 6: THE PRISON SURVEY 131
6.1 Sample Design 131
6.2 Gaining Cooperation 132
6.3 Interviewer Selection and Training 133
6.4 Data Collection 134
6.5 Quality Control 134

CHAPTER 7: PROCESSING THE DATA 136
7.1 Receipt Control 136
7.1.1 Screener 136
7.1.2 Background Questionnaire 137
7.1.3 Exercise Envelope 137
7.2 Coding and Scoring 138
7.2.1 Coding Background Questionnaires 138
7.2.2 Scoring Simulation Tasks 139
7.3 Data Entry 140
7.4 Editing and Quality Control 141

TABLE OF CONTENTS — CONTINUED

CHAPTER 8: ESTIMATING LITERACY PROFICIENCIES WITH AND WITHOUT COGNITIVE DATA 142
8.1 The Normal Treatment of Missing Cognitive Data 143
  8.1.1 Omitted Answers and Questions Not Reached 143
  8.1.2 Statistical Imputation Through Scaling 144
8.2 Reasons Cognitive Data Were Missing/Not Reached 146
  8.2.1 Non-interview Reports and Low Literacy Skills 151
  8.2.2 Internal Evidence for the Validity of Reasons 153
8.3 Using 'Reasons' to Improve Treatment of Missing Cognitive Data 155
  8.3.1 Five Logical Imputation Methods Considered 156
  8.3.2 Five Methods Applied to 1991 Field Test Data 158
  8.3.3 The Method Selected 163
8.4 Final Evaluation 164

CHAPTER 9: SCALING AND PROFICIENCY ESTIMATES 165
9.1 Scaling 165
9.2 Scaling Methodology 168
  9.2.1 The Scaling Model 168
  9.2.2 Design for Linking the 1992 Scales to the 1985 Scales 169
  9.2.3 Item Parameter Estimation 170
9.3 Proficiency Estimation Using Plausible Values 177
  9.3.1 Generating Proficiency Scores 177
  9.3.2 Linking the 1992 Scale to the 1985 Scale 182
  9.3.3 Evaluation of Differential Group Performance 183
9.4 Statistical Tests 187
  9.4.1 Analysis of Plausible Values 187
  9.4.2 Partitioning the Estimation Error Variance: A Numerical Example 188
  9.4.3 Minimum Sample Sizes for Reporting Subgroup Results 190
  9.4.4 Estimates of Standard Errors with Large Mean Squared Errors 190

CHAPTER 10: THE ROLE OF INCENTIVES IN LITERACY SURVEY RESEARCH 191
10.1 Literature Review 191
10.2 The 1991 Field Test 193
  10.2.1 Field Test Design 193
  10.2.2 Summary of Field Test Results 194
  10.2.3 Field Test Response Rates 196
  10.2.4 Representation of the Target Population in the Field Test 198
  10.2.5 Relationship Between Incentive Level, Self-Selection, and Performance in the Field Test 200
  10.2.6 Survey Costs for the Field Test 208
  10.2.7 Conclusions from the 1991 Field Test 209
10.3 The 1992 Incentive Experiment 210
  10.3.1 Sample Design for the 1992 National Adult Literacy Survey 210
  10.3.2 1992 Incentive Experiment Design 211
  10.3.3 Analysis of Response Rates 212
    10.3.3.1 Screener 213
    10.3.3.2 Background questionnaire and exercise booklet 214
10.4 Summary and Conclusion 217

TABLE OF CONTENTS — CONTINUED

CHAPTER 11: EVALUATION OF SAMPLE DESIGN AND COMPOSITE ESTIMATION 219
- 11.1 Estimation Methods 220
  - 11.1.1 Calculating Within- and Between-PSU Variance 222
- 11.2 The National Sample Design Evaluation 223
  - 11.2.1 Computing the Design Effect 228
- 11.3 Composite Estimation 237
  - 11.3.1 Estimating Compositing Factors Using the National Adult Literacy Survey Data 238
  - 11.3.2 Evaluating the National Adult Literacy Survey Compositing Factors 248
  - 11.3.3 Compositing Factors for a Future Assessment of Adult Literacy 248

CHAPTER 12: CONSTRUCT VALIDITY OF THE ADULT LITERACY SCALES 262
- 12.1 Data and Models 263
  - 12.1.1 Models for Four Restricted Adult Literacy Samples 263
  - 12.1.2 Models for Seven GED/Adult Literacy Samples 266
- 12.2 Method 269
- 12.3 Results and Discussion 271
  - 12.3.1 Restricted Adult Literacy Samples 271
  - 12.3.2 GED/Adult Literacy Samples 275
- 12.4 Conclusions 278

CHAPTER 13: INTERPRETING THE ADULT LITERACY SCALES AND LITERACY LEVELS 279
- 13.1 Background 279
- 13.2 The 1985 Young Adult Literacy Assessment 280
  - 13.2.1 Dimensionality of Literacy Skills 282
  - 13.2.2 Difficulty of Literacy Tasks 284
  - 13.2.3 Prose Comprehension Scale 285
  - 13.2.4 Document Literacy Scale 287
  - 13.2.5 Quantitative Literacy Scale 290
- 13.3 Enhancing Understanding of Task Difficulty 293
- 13.4 The 1990 Survey of the Literacy of Job-Seekers 296
  - 13.4.1 Prose Literacy 297
    - 13.4.1.1 Prose variables 297
    - 13.4.1.2 Examples of prose literacy tasks 299
    - 13.4.1.3 Coding the 1985 and 1990 prose literacy tasks 303
      - 13.4.1.3.1 Type of match 303
      - 13.4.1.3.2 Type of information 306
      - 13.4.1.3.3 Plausibility of Distractors 307
      - 13.4.1.3.4 Readability 308
    - 13.4.1.4 Codes for all 1985 and 1990 prose literacy tasks 308
    - 13.4.1.5 Validity evidence for the prose scale 310
  - 13.4.2 Documents Literacy 311
    - 13.4.2.1 Document variables 311
    - 13.4.2.2 Examples of document literacy tasks 312
    - 13.4.2.3 Coding the 1985 and 1990 document literacy tasks 314
      - 13.4.2.3.1 Type of match 314
      - 13.4.2.3.2 Plausibility of distractors 317
      - 13.4.2.3.3 Type of information 318
      - 13.4.2.3.4 Structural complexity 318
    - 13.4.2.4 Codes for all 1985 and 1990 document literacy tasks 320
    - 13.4.2.5 Validity evidence for the document scale 322

        13.4.3 Quantitative Literacy 323
            13.4.3.1 Quantitative variables 323
            13.4.3.2 Examples of quantitative literacy tasks 324
            13.4.3.3 Coding the 1985 and 1990 Quantitative Literacy Tasks 325
                13.4.3.3.1 Specificity of operation 327
                13.4.3.3.2 Type of calculation 328
                13.4.3.3.3 Type of match, plausibility of distractors, and structural complexity 329
            13.4.3.4 Codes for all 1985 and 1990 quantitative literacy tasks 329
            13.4.3.5 Validity evidence for the quantitative scale 330
        13.4.4 Establishing Proficiency Levels 331
    13.5 The 1992 National Adult Literacy Survey 336
        13.5.1 Prose Literacy 337
        13.5.2 Document Literacy 340
        13.5.3 Quantitative Literacy 343
        13.5.4 Levels of Prose, Document, and Quantitative Literacy 346
    13.6 Conclusion 347

CHAPTER 14: LITERACY LEVELS AND THE 80 PERCENT RESPONSE PROBABILITY CONVENTION 348
    14.1 Prose Literacy Tasks and Their Characteristics 349
    14.2 The Need for a Response Probability Convention 352
    14.3 Literacy Tasks and Literacy Levels 357
        14.3.1 Predictive Factors and RP80 Task Difficulty 358
        14.3.2 Predictive Factors, Task Difficulty, and the Response Probability Convention 360
        14.3.3 Alternative Cut Points between Literacy Levels 361
    14.4 Distribution of Adults Among Alternative Literacy Levels 365
    14.5 Selecting an Appropriate Response Probability Convention 367
    14.6 Conclusions 369

CHAPTER 15: WORKING WITH SPSS AND SAS 371
    15.1 The Electronic Code Book and SPSS and SAS Control Statements 371
    15.2 Creating SPSS System Files 372
    15.3 Creating SAS System Files 373
    15.4 Computing the Estimated Variance of a Statistic Using Jackknife Methods in SPSS or SAS 375

REFERENCES 380

APPENDIX A: ESTIMATED ITEM PARAMETERS A-1

APPENDIX B: CONDITIONING VARIABLES B-2

APPENDIX C: GAMMAS C-2

APPENDIX D: RP80S AND ITEM PROBABILITIES X 100 .... D-1

APPENDIX E: NON-INTERVIEW REPORT FORM E-1

APPENDIX F: FORM FOR INTERVIEWER'S OBSERVATION ON THE EXERCISE BOOKLET AND THE OBSERVATIONS F-1

APPENDIX G: ENGLISH BACKGROUND QUESTIONNAIRE FOR HOUSEHOLDS G-1

TABLE OF CONTENTS — CONTINUED

APPENDIX H: ENGLISH BACKGROUND QUESTIONNAIRE FOR PRISONS H-1

APPENDIX I: DERIVED VARIABLES I-1

APPENDIX J: SPECIAL CODES FOR CONTINUOUS VARIABLES J-1

APPENDIX K: COUNTRY OF BIRTH CODES K-1

APPENDIX L: NOTES ON SCORING L-1

APPENDIX M: NOTES ON VARIABLES M-1

APPENDIX N: IMPACT OF TREATMENT ON DISTRIBUTION OF SCALE SCORES N-1

APPENDIX O: ESTIMATED COMPOSITE FACTORS FOR A SELECTED SET OF DEMOGRAPHIC VARIABLES FOR ILLINOIS, INDIANA, LOUISIANA, NEW JERSEY, NEW YORK, OHIO, PENNSYLVANIA, TEXAS, AND WASHINGTON O-1

APPENDIX P: ELECTRONIC CODE BOOK FOR WINDOWS USER'S MANUAL P-1

APPENDIX Q: STANDARD ERRORS FOR TABLE 8.2 .... Q-1

APPENDIX R: STANDARD ERROR TABLES FOR CHAPTER 10 .... R-1

Chapter 1

THE NATIONAL ADULT LITERACY SURVEY: AN OVERVIEW

Lynn Jenkins, Wordsworth Writing and Editing (formerly with Educational Testing Service)
Stéphane Baldi, American Institutes for Research

1.1 INTRODUCTION

The Adult Education Amendments of 1988 required the U.S. Department of Education to submit a report to Congress defining literacy and measuring the nature and extent of literacy among adults in the nation. To satisfy these requirements, the National Center for Education Statistics (NCES) and the Division of Adult Education and Literacy planned a nationally representative household survey to assess the literacy skills of the adult population in the United States. In September 1989, NCES awarded a four-year contract for that purpose to Educational Testing Service (ETS) with a subcontract to Westat, Inc., for sampling and field operations.

The National Adult Literacy Survey is the third and largest assessment of adult literacy funded by the Federal government and conducted by ETS. The two previous efforts included a 1985 household survey of the literacy skills of 21- to 25-year-olds, funded by the U.S. Department of Education, and a 1989-90 survey of the literacy proficiencies of job seekers, funded by the U.S. Department of Labor.

In 1992, nearly 13,600 individuals age 16 and older, randomly selected to represent the adult population in this country, were surveyed in their homes. In addition, about 1,000 randomly selected adults age 16 through 65 were surveyed in each of 11 states that chose to participate in a concurrent State Adult Literacy Survey designed to produce state-level results comparable to the national data. In addition to the household samples, 1,147 inmates from 87 state and Federal prisons were randomly surveyed to represent the inmate population in the United States. Their participation helped to provide better estimates of the literacy levels of the total population and made it possible to report on the literacy proficiencies of this important segment of society.

Each individual who participated in the National and State Adult Literacy Surveys was asked to provide background demographic information and to complete a booklet of literacy tasks. These tasks were carefully constructed to measure respondents' ability to read and use a wide array of printed and written materials.

The survey results comprise an enormous set of data that includes more than a million responses to the literacy tasks and background questions. More important than the size of the database, however, is the fact that it provides information that is essential to understanding this nation's literacy resources. Specifically, the National Adult Literacy Survey data give policy makers, business and labor leaders, educators, researchers, and citizens vital information on the condition of literacy in the United States. The survey results can be used to:

- Describe the levels of literacy demonstrated by the adult population as a whole and by adults in various subgroups, including those targeted as "at risk;"
- Characterize adults' literacy skills in terms of demographic and background information (such as reading characteristics, education, and employment experiences);
- Profile the literacy skills of the nation's work force;
- Compare assessment results from the current study with those from the 1985 literacy survey of young adults;
- Interpret the findings in light of information-processing skills and strategies, so as to inform curriculum decisions concerning adult education and training; and
- Increase understanding of the skills and knowledge associated with living in a technological society.

This chapter describes the design for the 1992 National Adult Literacy Survey and gives an overview of the steps involved in its implementation, from the development of a working definition of literacy to the creation of edited data files. The major components of the implementation of the survey are presented here as a tool to help the reader gain an overview of the National Adult Literacy Survey without having to read each individual chapter. For more detailed or technical information, the reader is referred to the specific chapters of this technical report as well as to the booklet Assessing Literacy (Campbell, Kirsch, and Kolstad, 1992) and the initial report on the survey, Adult Literacy in America (Kirsch, Jungeblut, Jenkins, and Kolstad, 1993).

The organization of this chapter is as follows:

Section 1.2 provides an overview of the development of the working definition of literacy that underlies the National Adult Literacy Survey.

Section 1.3 summarizes the stratified random sampling procedures used for the national, state, and prison components of the survey.

Section 1.4 gives an overview of the use and computation of weights used in the 1992 National Adult Literacy Survey to permit inferences from persons included in the sample to the populations from which they were drawn.

Section 1.5 discusses the development of cognitive and background questions in the survey instrument.

Section 1.6 summarizes the field operations and data collection in the household and prison surveys.

Section 1.7 describes the data processing operations, including data entry, validation, the treatment of missing data, and the creation of edited data files.

Section 1.8 discusses the Item Response Theory (IRT) scaling model and the plausible values methodology used to score respondents' performance to the items in the questionnaire.

Section 1.9 discusses the establishment of literacy levels for the National Adult Literacy Survey.

1.2 DEFINING LITERACY

Although few would deny the importance of literacy in today's society, a shared belief in the value of literacy does not imply consensus on how to define and measure it. In fact, there are widely varying opinions about the skills that individuals need to function successfully in their work, in their personal lives, and in society, and about the ways in which these skills should be assessed. As a result, there have been widely conflicting diagnoses of the literacy problem in this country.

A committee of experts from business and industry, labor, government, research, and adult education worked with ETS staff to develop the definition of literacy that underlies the National Adult Literacy Survey, as well as to prepare the assessment objectives that guided the selection and construction of assessment tasks. In addition to this Literacy Definition Committee, a Technical Review Committee was formed to help ensure the soundness of the assessment design, the quality of the data collected, the integrity of the analyses conducted, and the appropriateness of the interpretations of the final results.

Drawing on the two earlier studies of adult literacy conducted by ETS and funded by the Federal government (Kirsch and Jungeblut, 1986; Kirsch, Jungeblut, and Campbell, 1992), the Literacy Definition Committee rejected the types of arbitrary standards—such as signing one's name, completing five years of school, or scoring at a particular grade level on a school-based measure of reading achievement—that have long been used to make judgments about adults' literacy skills. Through a consensus process, the committee adopted the following definition of literacy, initially developed for the 1985 young adult survey:

> *Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential.*

This definition of literacy extends beyond simple decoding and comprehension to include a broad range of skills that adults use in accomplishing many different types of literacy tasks associated with work, home, and community contexts.

1.3 THE SAMPLE

The National Adult Literacy Survey was administered to three samples: 1) a national household sample, 2) household samples from 11 states, and 3) a national sample of prison inmates. Both the national and state household samples were based on four-stage, stratified sampling. The prison sample was based on two-stage sampling. While the national and state household samples were drawn using the same sampling strategy, they differed in two ways: blacks and Hispanics were oversampled only in the national sample, and the target population for the national sample consisted of adults age 16 or older while for the state sample the target population consisted of adults ages 16 to 64. Blacks and Hispanics were oversampled in the national sample based on the key objective of the national sample: to provide reliable statistics for the adult population along with the prespecified domains. The prespecified domains included a racial/ethnic domain and an adults aged 65 and older domain. While the states wanted reliable statics, they were not concerned with the specific domains, and thus did not oversample them.

The four sampling stages for the national and state samples were: (1) the selection of primary sampling units (PSUs) consisting of counties or groups of counties, (2) the selection of segments consisting of census blocks or groups of blocks, (3) the selection of households, and (4) the selection of age-eligible individuals. In the first stage of sampling, the PSUs were stratified according to census region, metropolitan status, percentage of black residents, percentage of Hispanic residents, and, whenever possible, per capita income. In the second stage of sampling, census blocks or groups of blocks within each PSU were selected with a probability proportional to the number of housing units. In the third stage, a list of all housing units was created. A list of all housing units within the boundaries of each segment was then selected. Households were selected with equal probability within each segment of census blocks or groups of blocks, except for White, non-Hispanic households in high-minority segments in the national component. Finally, in the fourth stage of sampling, one person was randomly selected from each household with fewer than four eligible members and two persons were randomly selected from each household with four or more eligible members, from a list of all age-eligible household members (age 16 or older for the national sample and age 16 to 64 for the state samples). The same stratification methods, PSU construction, sample design and instruments were used for both the national and state designs.

In addition, at the request of the Office of Management and Budget, a subsample of 1,812 households drawn from the 2,064 segments in the national sample was randomly selected following the steps outlined above in order to yield approximately 1,000 respondents who would be administered the survey without a $20 incentive. This was done to be able to compare the incentive versus non-incentive response rates as well as assess the effect of incentives on response patterns.

For the prison survey, the two sampling stages were (1) the selection of primary sampling units (PSUs), and (2) the selection of inmates within each PSU. In this case, PSUs consisted of state or Federal adult correctional facilities, which were selected with a probability proportional to size. In the second stage, inmates were selected with a probability inversely proportional to the number of inmates, up to 22 inmates in a facility. Chapter 2 provides a discussion of the sample design.

1.4 WEIGHTING

Whenever various subsets of the population are sampled at different rates or have different rates of selection or response, weights are necessary in order to permit inferences from persons included in the sample to the populations from which they were drawn, as well as to have sample estimates reflect estimates of the larger population. For example, in the national component of the National Adult Literacy Survey, blacks and Hispanics were oversampled to ensure reliable estimates of literacy proficiencies and to permit analyses of the performance of different subpopulations. Furthermore, because only one person was selected in households with fewer than four eligible members, members of households with only one eligible member had twice the chance of selection as members of households with two eligible members, and three times the chance of selection as those in households with three eligible members. In such cases, weights are necessary to prevent serious bias in the estimates. Specifically, in the National Adult Literacy Survey, weights were computed to accomplish the following five objectives: (1) to permit unbiased estimates, taking account of the fact that all persons in the population did not have the same probability of selection, (2) to combine the state and national samples in an efficient manner, (3) to bring data up to the dimensions of the population totals, (4) to use auxiliary data on known population characteristics in such a way as to reduce sampling errors, and (5) to minimize biases arising from differences between cooperating and non-cooperating persons in the sample.

Differential probability of selection was corrected by computing base weights for all persons selected into the sample. For all three components (national, state, and prison), the base weight was calculated as the reciprocal of a respondent's final probability of selection. Furthermore, to combine the state and national samples, composite weights were calculated for the respondents in the 11 state samples and the respondents in the national sample PSUs in the 11 states. Finally, to adjust for non-response, weights were adjusted through post stratification and raking to match 1990 census totals. Chapter 3 provides detailed information on the weighting procedures.

1.5 THE SURVEY INSTRUMENT: MEASURING LITERACY

The Literacy Definition Committee endorsed the notion that literacy is neither a single skill suited to all types of texts, nor an infinite number of skills, each associated with a given type of text or material. Rather, as suggested by the results of the young adult and job seeker surveys, an ordered set of literacy skills appears to be called into play to accomplish diverse types of tasks. Accordingly, in addition to adopting the definition of literacy that guided the earlier young adult and job-seeker studies, the Literacy Definition Committee adopted three literacy scales—prose, document, and quantitative—to report the results of the surveys.

> *Prose literacy* involves the knowledge and skills needed to understand and use information from texts that include editorials, news stories, poems, and fiction; for example, finding a piece of information in a newspaper article, interpreting instructions from a warranty, inferring a theme from a poem, or contrasting views expressed in editorials.
>
> *Document literacy* concerns the knowledge and skills required to locate and use information contained in materials that include job applications, payroll forms, transportation schedules, maps, tables, and graphs; for example, locating a particular intersection on a street map, using a schedule to choose the appropriate bus, or entering information on an application form.
>
> *Quantitative literacy* involves the knowledge and skills required to apply arithmetic operations, either alone or sequentially, using numbers embedded in printed materials; for example, balancing a

checkbook, figuring out a tip, completing an order form, or determining the amount of interest from a loan advertisement.

The prose, document, and quantitative scales were augmented in the current survey through the addition of new assessment tasks that took into account the following:

- Continued use of open-ended simulation tasks;
- Continued emphasis on tasks that measure a broad range of information-processing skills and cover a wide variety of contexts;
- Increased emphasis on simulation tasks that require brief written and/or oral responses;
- Increased emphasis on tasks that ask respondents to describe how they would set up and solve a problem; and
- Use of a simple, four-function calculator to solve selected quantitative problems.

Approximately 110 new assessment tasks were field tested, and 81 of these were selected for inclusion in the survey. These 81 new assessment tasks were added to a pool of 85 tasks that were administered in both the young adult and job-seeker assessments (Kirsch and Jungeblut, 1986a and 1992). Thus, the National Adult Literacy Survey consisted of a total of 166 assessment tasks. By administering a common set of assessment tasks in each of the three literacy surveys, it is possible to compare results across time and across population groups.

No individual could be expected to respond to the entire set of 166 simulation tasks administered as part of the National Adult Literacy Survey. It was therefore necessary to adopt a survey design that would give each person participating in the study a subset of the total pool of literacy tasks, while at the same time ensuring that each of the 166 tasks was administered to a nationally representative sample of the adult population. Literacy tasks were assigned to blocks or sections that could be completed in about 15 minutes, and these blocks were then compiled into booklets in such a way that each block appeared in each position (first, middle, and last) and each block was paired with every other block. Thirteen blocks of simulation tasks were assembled into 26 booklets, each of which could be completed in about 45 minutes. During a personal interview, each survey participant was asked to complete one booklet.

In addition to the time allocated for the literacy tasks, approximately 20 minutes were devoted to obtaining personal information from respondents. Major areas explored included background demographics, education, labor market experiences, income, and literacy-related activities. These background data help to improve understanding of the ways in which various characteristics are associated with demonstrated literacy skills.

Trained interviewers surveyed some 13,600 adults age 16 and older, chosen to represent the household population nationwide. In addition to the national samples, approximately 1,000 adults ages 16 to 64 were assessed in each of the states that chose to participate in the State Adult Literacy Survey, a special study designed to provide state-level data comparable to the national results. California, Illinois, Indiana, Iowa, Louisiana, New Jersey, New York, Ohio, Pennsylvania, Texas, and Washington conducted their surveys at the same time as the national survey. (One additional state, Florida, was surveyed at a later date.) To permit comparisons of the state and national results, the survey instruments administered to the state and national samples were identical.

Finally, 1,147 inmates from 87 state and Federal prisons were surveyed. Because some questions included in the household survey were inappropriate for the prison population, a revised version of the background questionnaire was developed that included queries about current offenses, criminal history, and prison work assignments, as well as education and work force experiences. To ensure comparability with the national survey, the simulation tasks (tasks that simulate the demands that adults encounter when they interact with printed materials on a daily basis) given to the prison participants were the same as those given to the household survey population.

A total of 26,091 adults gave, on average, over an hour of their time to complete the National Adult Literacy Survey instruments. Those who agreed to participate in the survey and completed as much of the assessment as their skills allowed were paid $20 for their time. Responses from the national, state, and prison samples were combined to yield the best possible performance estimates. Chapter 4 describes the development of the survey instrument.

1.6 FIELD OPERATIONS

Field operations and data collection for the National Adult Literacy Survey were the responsibility of Westat, Inc. The literacy survey was conducted between February and August 1992 by more than 400 trained interviewers, some of whom were bilingual in English and Spanish. All components of the survey sample were worked simultaneously, including the national sample, the state sample, and the prison sample. The field organization was headed by the survey field director, who reported directly to the Westat project director and who was supported by four home-office field managers and 24 field supervisors located across the United States. Each supervisor was supported in the field by an editor who was responsible for completely editing each case received from the field.

Interviewers were recruited directly based on Westat's computerized field personnel file containing information on over 4,000 field staff who had worked for Westat in the previous three years. A total of 456 interviewers were recruited, of which 2 did not attend training and 2 were released at training. Training consisted of a 3-day in-person training program, preceded by home study.

The administration of the national and state household surveys to respondents occurred in three overlapping stages: an initial phase, in which each area segment was assigned to an interviewer; a reassignment phase, in which incomplete interviews were given to another interviewer in the same PSU; and a special non-response conversion phase, in which the home office assembled a special traveling team of the most experienced interviewers to perform a non-response conversion effort.

For the survey of the prison population, 51 interviewers were recruited from among the household survey workforce. These interviewers received an additional 1-day, in-person training session emphasizing collecting data on criminal history and prison employment. Interviewers were required to perform a careful edit before leaving the facility because it was not possible to recontact the prisoners if errors were made.

An automated management system tracked and recorded the progress of fieldwork throughout the interview phase. In addition, progress was monitored weekly through telephone conferences between field supervisors, Westat home office staff, and ETS staff. Quality control checks were performed throughout the field data collection period and took the form of careful editing of completed documents, validation of 10 percent of each interviewer's closed-out cases, observations of interviews in person and by tape recordings, and observation of supervisors by the Westat home office and ETS staff.

As a result of the careful design of the field operations, the response rates achieved were quite favorable. Eighty-one percent of eligible respondents for the combined state and national surveys answered the background questionnaire. Of those, 95.8 percent completed the booklet of literacy exercises. For the prison population, 85.6 percent completed the background questionnaire, and 96.2 percent of those completing the background questionnaire completed the exercise booklet. Chapters 5 and 6 document the field operations for the household and prison surveys respectively.

1.7 DATA PROCESSING AND MISSING DATA

After performing quality checks on completed background questionnaires and exercise booklets, field supervisors shipped them to ETS where staff checked the contents of each shipment against the enclosed transmittal form serving as the packing list for the shipment. The background questionnaires were then given to coders who coded the open-ended items, and the exercise booklets were given to readers who scored the open-ended literacy items.

Coding was performed by 20 individuals, 9 working on the background questionnaire and 11 on the exercise booklets, following coding guides developed by scoring supervisors. To check the accuracy of coding in the background questionnaire, items dealing with country of birth, language, wages, and date of birth were checked in 10 percent of the questionnaires by a second coder. In the exercise booklets, 20 percent of all booklets were checked by a second coder who performed a reliability check. The inter-reader reliability for booklets scored by two readers was 97 percent, a number comparing very favorably with the reliability for the 1985 young adult literacy assessment.

The coded responses for the background questionnaire and exercise booklets were then recorded onto scannable answer sheets that were then scanned by ETS staff and transmitted to magnetic tape. The data were then transferred to a database on the main computer for editing and quality control. In a final stage, the data files were examined for nonexistent housing locations, illogical or inconsistent responses, multiple responses, as well as to insure that the skip patterns had been properly followed and that all data errors had been resolved.

In order to address the issue of missing data, several imputation methods were considered using field test data as well as non-interview report data collected by the interviewers. Three of the five imputation methods made no use of the non-interview report data and the remaining two were informed by the reasons found in the non-interview report. A series of analyses examined the extent to which using each of the five imputation methods affected overall literacy proficiency estimates. Because imputation methods which made no use of the non-interview report data tended to weaken the educational, income, and racial/ethnic differences in literacy scores, they were ruled out, leaving two viable imputation methods. After consulting with others and examining the analyses performed using the two remaining imputation methods, the Technical Review Committee and the Literacy Definition Committee advising the National Adult Literacy Survey project adopted an imputation method for dealing with missing responses. When a respondent failed to answer consecutive assessment tasks and cited a reason related to literacy skills (e.g., "I can't read these tasks"), the missing tasks were assigned wrong answers. That is, they were scored as if the respondent had attempted and failed the tasks. The extensive processing of the data is detailed in Chapter 7. Chapter 8 provides a discussion of the missing data procedures.

1.8 SCALING AND PROFICIENCY ESTIMATES

The scaling model used for the National Adult Literacy Survey is the three-parameter (3PL) model from item response theory (Birnbaum, 1968; Lord, 1980). This model estimates the probability that an individual will respond correctly to a particular task from a single domain of tasks as a function of a parameter characterizing the proficiency of that individual and three parameters characterizing the properties of a given task in terms of its sensitivity to proficiency, its difficulty, and its non-zero chance of correct response for a multiple-choice task. Item response theory (IRT) models are based on the assumptions of conditional independence (i.e., item response probabilities depend only on a measure of proficiency and the specified item parameters) and unidimensionality (i.e., performance on a set of items is accounted for by a single variable). Thus, a critical part of the data analysis involved the testing of these two assumptions in order to validate the accuracy and integrity of the results.

Because in the National Adult Literacy Survey each respondent was administered relatively few items in a subject area scale, comparing scale scores based on the respondents' responses to different questions would lead to seriously biased estimates of proficiency. To circumvent this problem, proficiency scores for respondents were estimated using plausible values methodology. Plausible values provide consistent estimates of population characteristics, even though they are not unbiased estimates of the proficiencies of the individuals with whom they are associated. Thus, plausible values are not test scores for individuals in the usual sense. They are merely an intermediate measure used to estimate population characteristics. Chapter 9 discusses the scaling methodology as well as the calculation of proficiency estimates using plausible values methodology (Mislevy, Beaton, Kaplan, Sheehan, 1993).

1.9 ESTABLISHING LITERACY LEVELS

As previously noted, the results of the National Adult Literacy Survey are reported using three scales: a prose scale, a document scale, and a quantitative scale. The literacy scales, each ranging from 0 to 500, provide a useful way to describe the various types and levels of literacy demonstrated by adults in the population as a whole and in different subpopulations. The scales used an item mapping procedure reflecting response probabilities (RP). Tasks were placed on the scale at the point at which a minimum of 80 percent (i.e., RP80) of respondents at a particular ability level could be expected to complete the task successfully. The scores on each literacy scale represent degrees of proficiency along that particular dimension of literacy. For example, a low score (below 200) on the document scale indicates that an individual has very limited skills in processing information from tables, charts, graphs, maps, and the like (even those that are brief and uncomplicated). On the other hand, a high score (above 375) indicates advanced skills in performing a variety of tasks that involve the use of complex documents.

The literacy scales also make it possible to determine the relative difficulty of the literacy tasks included in the survey. In other words, just as individuals receive scale scores according to their performance in the assessment, the literacy tasks receive different scale values according to their difficulty, as determined by the performance of the adults who participated in the survey. The literacy tasks administered in the National Adult Literacy Survey varied widely in terms of materials, content, and task requirements, and thus in difficulty. A careful analysis of the range of tasks along each scale provides clear evidence of an ordered set of information-processing skills and strategies along each scale. To capture this ordering, each scale was divided into five levels that reflect this progression of information-processing skills and strategies: Level 1 (0 to 225), Level 2 (226 to 275), Level 3 (276 to 325), Level 4 (326 to 375), and Level 5 (376 to 500). By examining the tasks within each literacy level, it is possible to identify the types of materials or directives that are more or less difficult for various types of readers. Further, by examining the characteristics of individuals who performed at each literacy level, it is possible to identify factors associated with higher or lower proficiency in reading and using prose, documents, or quantitative materials. Chapter 13 summarizes the establishment of literacy levels for the National Adult Literacy Survey.

Appendices A through R, respectively, contain information about: estimated item parameters, conditioning variables, gamma values, RP80s and item probabilities, the non-interview report form, interviewer's observation guide, English background questionnaire for households, English background questionnaire for prisons, derived variables, codes for continuous variables, birth codes, scoring the variables, sample-specific variables, treatment distribution, estimated composite factors, the code book for windows, and standard errors for Chapters 8 and 10.

Chapter 2

SAMPLE DESIGN

Leyla Mohadjer, Joseph Waksberg, Huseyin Goksel, and James Green, Westat, Inc.

2.1 OVERVIEW

The National Adult Literacy Survey included the following three components: 1) a national household sample; 2) household samples from 11 states; and 3) a national sample of prison inmates.

The national and state household components were based on a four-stage, stratified area sample with the following stages: (1) the selection of primary sampling units (PSUs) consisting of counties or groups of counties, (2) the selection of segments consisting of census blocks or groups of blocks, (3) the selection of households, and (4) the selection of age-eligible individuals. A single area sample was drawn for the national component, and 11 additional state-level area samples were drawn for the state component (i.e., California, Illinois, Indiana, Iowa, Louisiana, New Jersey, New York, Ohio, Pennsylvania, Texas, and Washington).[1] The national and state samples differed in two important respects. In the national sample, Black and Hispanic individuals were sampled at a higher rate than the remainder of the population to increase their representation in the sample, whereas the state samples used no oversampling. Also, the target population for the national sample consisted of adults age 16 or older, whereas the target population for the state samples consisted of adults ages 16–64.

As noted above, the first stage of sampling for all 12 household samples involved the selection of PSUs, which consist of counties or groups of counties. The PSUs were stratified according to census region, metropolitan status, percentage of Black residents, percentage of Hispanic residents, and, whenever possible, per capita income. The national component used a 101-PSU sample. The national frame of PSUs was used to construct individual state frames for the state components, and a sample of 8 to 12 PSUs was selected within each of the 11 states. All PSUs were selected with a probability proportional to the PSUs' 1990 population.

For the second stage of sampling, segments (census blocks or groups of blocks) within the PSUs were selected with a probability proportional to size, where the measure of size for a segment was a function of the number of year-round housing units within the segment. The oversampling of Black and Hispanic persons for the national component was carried out at the segment level, where segments were classified as high minority (segments with more than 25 percent Black or Hispanic residents) or low minority. The measure of size for high-minority segments was defined as the number of White, non-Hispanic households plus three times the number of Black or Hispanic households. High-minority

---

[1] A state-level survey was later conducted in Florida, but the data are not included in this report.

segments were therefore oversampled at up to three times the rate of low-minority segments. As for all segments in the state components, the measure of size was simply the number of year-round housing units within the segment. One in seven of the national component segments was selected at random to be included in a "non-incentive" sample (see section 2.3 for more details). Respondents from the remaining segments in the national component received a monetary incentive for participation, as did all respondents in the state components. Data for respondents from the non-incentive segments were not included in the analyses reported by the National Center for Education Statistics, but are available as one of the three principal analysis files (household, prison, and non-incentive data).

For the third stage of sampling, the selection of households within segments, Westat field staff visited all selected segments and prepared lists of all housing units within the boundaries of each segment, as determined by the 1990 census block maps. The lists were used to construct the sampling frame for households. Households were selected with equal probability within each segment, except for White, non-Hispanic households in high-minority segments in the national component. These households were sub-sampled after screening, so that the sampling rates for White, non-Hispanic persons would be about the same in the high-minority segments as in other segments.

For the fourth stage of sampling, a list of age-eligible household members (age 16 or older for the national component, 16–64 for the state component) was constructed for each selected household. One person was selected at random from households with fewer than four eligible members, and two persons were selected at random from households with four or more eligible members. The interviewers were instructed to list the eligible household members in descending order of age. The interviewers then identified the one or two sample household members based on computer-generated sampling messages that had been attached to each questionnaire in advance.

The sample design for the prison component involved two stages of selection. For the first stage of sampling, state or Federal correctional facilities were selected with a probability proportional to size, where the measure of size for a facility was equal to the size of the inmate population. The second stage involved the selection of inmates within each facility. Inmates were selected with a probability inversely proportional to the size of their facility's inmate population (up to 22 inmates in a facility). Table 2-1 provides the sample sizes for all stages of sampling for the national and state components of the National Adult Literacy Survey.

Section 2.2 provides a review of the four stages of sampling for the national component of the survey. A similar discussion of the state samples is presented in section 2.4. Section 2.5 presents weighted and unweighted response rates for the household component of the survey. Sections 2.3 and 2.6 describe the non-incentive sample design and the prison sample design, respectively.

Table 2-1. Sample sizes for the national and state components of the National Adult Literacy Survey

Component	Number of PSUs	Number of segments•	Number of households*	Number of persons screened	Number of persons interviewed	Number of persons assessed
National and state incentive sample	210	3,733	43,783	30,806	24,944	22,107
National non-incentive sample	101	155	1,838	1,273	930	695
State samples**						
CA	20	405	4,917	3,371	2,665	2,143
IL	14	262	2,914	2,130	1,668	1,504
IN	15	215	2,361	1,755	1,441	1,368
IA	14	187	2,041	1,446	1,246	1,192
LA	10	188	2,270	1,460	1,192	1,087
NJ	16	243	2,790	1,821	1,317	1,111
NY	14	302	3,526	2,139	1,688	1,415
OH	17	246	2,691	1,984	1,568	1,510
PA	14	253	2,950	2,060	1,626	1,532
TX	16	316	3,833	2,681	2,209	1,834
WA	9	182	2,096	1,506	1,244	1,186

• The numbers include segments with at least one dwelling unit selected into the sample.

* The numbers include the missed structures and units (refer to section 2.2.3.3) incorporated into the sample during the data collection.

** Numbers include the national sample cases in each state in addition to the individually selected state sample.

2.2 SAMPLING FOR THE NATIONAL COMPONENT

The target population for the national component of the National Adult Literacy Survey consisted of adults age 16 or older in the 50 states and the District of Columbia who, at the time of the survey (February through August, 1992), resided in private households or college dormitories.

The household component used a four-stage, stratified sample design. The first-stage sample was a sample of PSUs (counties or groups of counties) developed by Westat. In developing the sampling frame, the 3,141 counties and independent cities in the 50 states were grouped into 1,404 PSUs, from which a sample of 101 PSUs was selected for the household component. In the second stage of sampling, probability sampling was used to select a sample of 2,064 segments (census blocks or combinations of blocks) from the PSUs chosen during the first stage. The third stage of sampling involved the selection of 24,522 housing units from listings developed within the selected segments by the field listers. In the fourth stage, age-eligible persons were chosen for interview and assessment from within selected households. The stages of sampling for the national component are described in greater detail in the following sections.

2.2.1 First-Stage Sample

The first-stage sample was a sample of PSUs (counties or groups of counties) developed by Westat.

2.2.1.1 Westat's master sample of PSUs

In selecting the master sample, Westat used the 1990 census Public Law 94-171 (PL94) data tape file as the source of information (total and minority population sizes for each county) for stratification as well as to determine PSU size. The income data were based on the 1988 per capita income reported by the Bureau of Economic Analysis.

In designing the Westat PSU sample, entire metropolitan statistical areas (MSAs) were treated as single PSUs; however, because of their size, the New York, Los Angeles, and Chicago MSAs were divided into three, two, and two PSUs, respectively. In New England, whole-county approximations of MSAs were used. Counties outside of MSAs were grouped to make PSUs (1) large enough to provide a sufficient sample size for most national surveys and (2) as internally heterogeneous as possible but still small enough that an interviewer could conveniently travel across the PSU. A total of 1,404 PSUs were constructed. All PSUs consisted of one or more contiguous counties, or contiguous counties and independent cities, and had minimum population sizes of 15,000. Additionally, all PSUs were completely contained within the boundaries of one of the four census regions.

Master sample PSUs were stratified on the basis of the social and economic characteristics of the population, as reported in the 1990 census. Strata were of roughly equal size; they did not cross regions, and a stratum did not include both metropolitan and non-metropolitan PSUs. The following characteristics were used in stratifying the Westat PSUs (some explicitly and some implicitly, by ordering the PSUs and sampling systematically):

- Region of the country (four census regions);
- Whether or not the PSU was an MSA;
- Percentage of Black residents;
- Percentage of Hispanic residents; and
- Average income.

2.2.1.2 Selecting the sample of PSUs for the national component

The sampling frame for the Westat PSU sample included Hawaii and Alaska, but neither of the Hawaii or Alaska counties were selected for the 100-PSU master sample. Honolulu MSA was added to the sample as the 101st PSU in the national sample. Westat adjusted the weights to correctly account for the inclusion of the Honolulu PSU in the sample.

Table 2-2 shows the distribution of the population in the 101 PSUs selected for the household component of the survey. The measure of size for each PSU was equal to the 1990 population of the PSU.

Twenty-five PSUs were included in the sample with certainty on the basis of their sizes. Then 38 strata of approximately equal size were formed. Two PSUs were selected (without replacement), with probability proportionate to size, from each of the 38 strata. Among the multiple-PSU strata, 26 were MSA strata and 12 were non-MSA strata.

Table 2-2. Proportion of U.S. population in PSUs selected for the national component by stratum type, total 1990 population, Black, and Hispanic

		Total 1990		Black		Hispanic	
Stratum type	PSU sample	Number	%	Number	%	Number	%
Certainty MSA	**Total in frame**	**76,349,843**	**30.7**	**12,304,548**	**40.0**	**11,769,950**	**52.7**
Non-certainty MSA	**Total in frame**	**116,764,722**	**47.0**	**12,823,091**	**42.8**	**8,444,362**	**37.8**
	Not in sample	75,474,068	30.4	8,115,899	27.1	5,387,275	24.1
	In 101-PSU sample	41,290,654	16.6	4,707,192	15.7	3,057,087	13.7
Non-certainty non-MSA	**Total in frame**	**55,595,308**	**22.4**	**4,858,421**	**16.2**	**2,139,747**	**9.6**
	Not in sample	54,058,657	21.7	4,742,122	15.8	2,072,580	9.3
	In 101-PSU sample	1,536,651	0.6	116,299	0.4	67,167	0.3
	Grand total	**248,709,873**	**100.0**	**29,986,060**	**100.0**	**22,354,059**	**100.0**

Table 2-3 contains a listing of the 101 PSUs in the national sample (certainty PSUs are in bold).

Table 2-3. National Adult Literacy Survey 101-PSU sample

PSU	County and State
101	**Boston, MA**
	Essex
	Middlesex
	Norfolk
	Plymouth
	Suffolk
102	Pittsfield, MA
	Berkshire
103	Springfield, MA
	Hampden
	Hampshire
104	Providence, RI
	Bristol
	Kent
	Providence
	Washington
105	Newport, RI
	Newport
106	**Nassau/Suffolk, NY**
	Nassau
	Suffolk
107	**Kings/Richmond, NY**
	Kings
	Richmond
108	**New York/Queens, NY**
	New York
	Queens
109	**Bronx/Putnam, NY**
	Bronx
	Putnam
	Rockland
	Westchester
110	Rochester, NY
	Livingston
	Monroe
	Ontario
	Orleans
	Wayne
111	Buffalo, NY
	Erie
112	Bergen/Passaic, NJ
	Bergen
	Passaic
113	Newark, NJ
	Essex
	Morris
	Sussex
	Union
114	Monmouth/Ocean, NJ
	Monmouth
	Ocean
115	Atlantic City,NJ
	Atlantic
	Cape May
116	**Philadelphia, PA/Camden, NJ**
	Burlington, NJ
	Camden, NJ
	Gloucester, NJ
	Bucks, PA
	Chester, PA
	Delaware, PA
	Montgomery, PA
	Philadelphia, PA
117	Scranton/Wilkes-Barre, PA
	Columbia
	Lackawanna
	Luzerne
	Monroe
	Wyoming
118	Harrisburg, PA
	Cumberland
	Dauphin
	Lebanon
	Perry
119	Pittsburgh, PA
	Allegheny
	Fayette
	Washington
	Westmoreland

Certainty PSUs are in bold.

Table 2-3. 101-PSU sample – Continued

PSU	County and State
120	Butler, PA
	Butler
	Lawrence
201	Steubenville, OH (Weirton,WV)
	Jefferson
202	Youngstown/Warren, OH
	Mahoning
	Trumbull
203	Akron, OH
	Portage
	Summit
204	Cleveland, OH
	Cuyahoga
	Geauga
	Lake
	Medina
205	Cincinnati, OH/Dearborn, IN (Covington, KY)
	Dearborn, IN
	Clermont, OH
	Hamilton, OH
	Warren, OH
206	Saginaw/Bay City/Midland, MI
	Bay
	Midland
	Saginaw
207	**Detroit, MI**
	Lapeer
	Livingston
	Macomb
	Monroe
	Oakland
	St. Clair
	Wayne
208	Fountain/Montgomery/Putnam,IN
	Fountain
	Montomery
	Putnam
209	Indianapolis, IN
	Boone
	Hamilton
	Hancock
	Hendricks
	Johnson
	Marion
	Morgan
	Shelby
210	Gary/Hammond, IN
	Lake
	Porter
211	**Chicago, IL (CITY)**
	Chicago City
212	**Cook/DuPage/McHenry, IL (Chicago)**
	Cook
	DuPage
	McHenry
213	Aurora/Elgin, IL
	Kane
	Kendall
214	Knox/Mercer, IL
	Knox
	Mercer
215	Peoria, IL
	Peoria
	Tazewell
	Woodford
216	**St. Louis, MO/E. St. Louis, IL**
	Clinton, IL
	Jersey, IL
	Madison, IL
	Monroe, IL
	St.Clair, IL
	Franklin, MO
	Jefferson, MO
	St. Charles, MO
	St. Louis City, MO
	St. Louis, MO

Certainty PSUs are in bold.

Table 2-3. 101-PSU sample – Continued

PSU	County and State
217	Pike/Ralls, MO
	Pike
	Ralls
218	Howard/Saline, MO
	Howard
	Saline
219	Milwaukee, WI
	Milwaukee
	Ozaukee
	Washington
	Waukesha
220	**Minneapolis/St. Paul, MN/WI**
	Anoka, MN
	Carver, MN
	Chisago, MN
	Dakota, MN
	Hennepin, MN
	Isanti, MN
	Ramsey, MN
	Scott, MN
	St.Croix, MN
	Washington, MN
	Wright, WI
221	Iowa City, IA
	Johnson
222	Monona, IA/Thurston, NE
	Monona, IA
	Thurston, NE
223	Hall/Hamilton, NE
	Hall
	Hamilton
224	Cheyenne/Rooks, KS
	Cheyenne
	Decatur
	Graham
	Rawlins
	Rooks
	Sheridan
225	Atchison/Jackson/Jefferson, KS
	Atchison
	Jackson
	Jefferson
301	**Washington, D.C./MD/VA**
	District of Columbia
	Calvert, MD
	Charles, MD
	Frederick, MD
	Montgomery, MD
	PrinceGeorges, MD
	Arlington, VA
	Fairfax, VA
	Loudoun, VA
	PrinceWilliam, VA
	Stafford, VA
	Alexandria City, VA
	Fairfax City, VA
	Falls Church City, VA
	Manassas, VA
	Manassas Park, VA
302	Wilminton, DE/Cecil, MD
	New Castle, DE
	Cecil, MD
303	**Baltimore,MD**
	AnneArundel
	Baltimore County
	Baltimore City
	Carroll
	Harford
	Howard
	Queen Annes
304	Weirton, WV (Steubenville, OH)
	Brooke
	Hancock
305	Charlottesville, VA
	Albemarle
	Fluvanna
	Greene
	Charlottesville City
306	Norfolk/Virginia Beach,VA
	Gloucester
	James City
	York

Certainty PSUs are in bold.

Table 2-3. 101-PSU sample – Continued

PSU	County and State
	Chesapeake City
	Hampton City
	Newport News City
	Norfolk City
	Poquoson
	Portsmouth City
	Suffolk City
	Virginia Beach City
	Williamsburg City
307	Johnson City, TN/Bristol, VA
	Carter, TN
	Hawkins, TN
	Sullivan, TN
	Unicoi, TN
	Washington, TN
	Scott, VA
	Washington, VA
	Bristol City, VA
308	Covington, KY (Cincinnati, OH)
	Boone
	Campbell
	Kenton
309	Fort Knox, KY
	Breckinridge
	Grayson
	Meade
310	Greensboro/Winston-Salem NC
	Davidson
	Davie
	Forsyth
	Guilford
	Randolph
	Stokes
	Yadkin
311	Albemarle, NC
	Montgomery
	Stanly
312	Fayetteville, NC
	Cumberland
313	Nashville, TN
	Cheatham
	Davidson
	Dickson
	Robertson
	Rutherford
	Sumner
	Williamson
	Wilson
314	Chattanooga, TN/Dade, GA
	Catoosa, GA
	Dade, GA
	Walker, GA
	Hamilton, TN
	Marion, TN
	Sequatchie, TN
315	**Atlanta, GA**
	Barrow
	Butts
	Cherokee
	Clayton
	Cobb
	Coweta
	DeKalb
	Douglas
	Fayette
	Forsyth
	Fulton
	Gwinnett
	Henry
	Newton
	Paulding
	Rockdale
	Spaulding
	Walton
316	Greene/Lincoln, GA
	Greene
	Lincoln
	Oglethorpe
	Wilkes
317	Wheeler/Toombs, GA
	Montgomery
	Toombs

Certainty PSUs are in bold.

Table 2-3. 101-PSU sample – Continued

PSU	County and State
	Treutlen
	Wheeler
318	Tallahassee, FL
	Gadsden
	Leon
319	Tampa/St. Petersburg, FL
	Hernando
	Hillsborough
	Pasco
	Pinellas
320	Orlando, FL
	Orange
	Osceola
	Seminole
321	**Miami/Ft. Lauderdale, FL**
	Broward
	Dade
322	Birminham, AL
	Blount
	Jefferson
	St. Clair
	Shelby
	Walker
323	Dothan, AL
	Dale
	Houston
324	Meridian, MS
	Lauderdale
	Newton
325	Franklin/Madison, AR
	Franklin
	Madison
326	Pope, AR
	Pope
327	Shreveport, LA
	Bossier
	Caddo
328	Muskogee/McIntosh, OK
	McIntosh
	Muskogee
329	**Dallas, TX**
	Collin
	Dallas
	Denton
	Ellis
	Kaufman
	Rockwall
330	Anderson TX
	Anderson
331	Austin, TX
	Hays
	Travis
	Williamson
332	San Antonio, TX
	Bexar
	Comal
	Guadalupe
333	**Houston, TX**
	Fort Bend
	Harris
	Liberty
	Montgomery
	Waller
334	Big Spring, TX
	Howard
401	Seattle, WA
	King
	Snohomish
402	Portland, OR
	Clackamas
	Multnomah
	Washington
	Yamhill
403	Missoula, MT
	Missoula

Certainty PSUs are in bold.

Table 2-3. 101-PSU sample – Continued

PSU	County and State
404	Boise City, ID
	Ada
405	Elmore/Twin Falls, ID
	Elmore
	Twin Falls
406	Sacramento, CA
	El Dorado
	Placer
	Sacramento
	Yolo
407	**San Francisco/Oakland, CA**
	Alameda
	Contra Costa
	Marin
	San Francisco
	San Mateo
408	San Jose, CA
	Santa Clara
409	Merced, CA
	Merced
410	Fresno, CA
	Fresno
411	**Riverside/San Bernardino, CA**
	Riverside
	San Bernardino
412	**Los Angeles City, CA**
	Los Angeles City
413	**Los Angeles/Long Beach, CA**
	Los Angeles
414	**Anaheim/Santa Ana, CA**
	Orange
415	**San Diego, CA**
	San Diego
416	Douglas/Storey/Carson City, NV
	Douglas
	Storey
	Carson City
417	Las Vegas, NV
	Clark
418	**Phoenix, AZ**
	Maricopa
419	Tucson, AZ
	Pima
420	Cibola/Valencia, NM
	Cibola
	Valencia
421	Boulder, CO
	Boulder
422	Honolulu, HI

Certainty PSUs are in bold.

2.2.2 Second-Stage Sample—Selecting Census Blocks (Segments)

Within each PSU, area segments consisting of census blocks (or combinations of two or more adjacent census blocks) were selected with probability proportionate to size. A total of 2,064 segments were chosen, an average of 21 per PSU. The frame for defining and sampling segments was the 1990 PL94 data.

The sample design requirements called for an average cluster size of about seven interviews (i.e., an average of about seven completed background interviews per segment). The sample of housing units within each segment was designed to account for attrition. Attrition was expected because, according to figures obtained from the 1990 census, approximately 10 percent of the housing units were probably vacant. Additionally, we expected a 10 percent screener refusal rate and a 15 percent background questionnaire refusal rate. The sample of housing units selected within each segment was thus made equal to 11. In addition, a reserve sample of approximately 5 percent of the size of the main sample was selected and set aside in case of shortfalls due to unexpectedly high vacancy and nonresponse rates.

2.2.2.1 Measures of size and sampling rates

Standard texts on sampling discuss measure of size in multistage designs for household surveys only in univariate situations. In effect, they describe how the total population can be used as the measure of size when sampling areas with probability proportionate to size, followed by sampling within each area at a rate proportionate to the reciprocal of the measure of size. A sample selected in this way has two desirable properties: (1) it is a self-weighting sample (i.e., all households are selected at the same rate), and (2) the interviewer workloads are approximately the same in all areas. The second property provides operational efficiency and results in lower variances than designs with variable workloads.

The national sample design modified and adapted the theory for multivariate situations by establishing a measure of size that produced constant workloads among segments and, at the same time, produced constant (but separate) sampling rates for minorities and non-minorities within each of two strata. The following is a description of the derivation of measures of size for this survey.

One of the requirements of the national design was to sample Black and Hispanic adults at a higher rate than the remainder of the population. Segments where 25 percent or more of the population consisted of Black and Hispanic adults were oversampled at a rate up to three times that of the remainder of the segments.

The housing unit counts served as the measure of size for the low-minority segments (segments with less than 25 percent Black or Hispanic households). In high-minority segments, the measure of size was equal to the number of White, non-Hispanic households plus three times the number of Black and Hispanic households. In low-minority segments, the measure of size of a segment was equal to the number of households in the segment.

$$MOS_{2ij} = H_{Oij} + H_{Mij} \qquad (1)$$

where

MOS_{2ij} = measure of size for the ij^th segment in the low-minority stratum.

H_{Oij} = number of "other" (i.e., non-minority) households in the j^th segment in the i^th PSU; and

H_{Mij} = number of minority (Black plus Hispanic) households in the j^th segment in the i^th PSU;

In high-minority segments, the measure of size was equal to

$$MOS_{1ij} = H_{Oij} + 3H_{Mij}.$$

where

MOS_{1ij} = measure of size of the ijth segment in the high-minority stratum (the minority stratum is defined as segments in which the Black plus Hispanic population is 25 percent or more of the total population).

The sampling interval, I, was computed as

$$I = \frac{\sum_{1_{ij}} MOS_{1ij}/P_i + \sum_{2_{ij}} MOS_{2ij}/P_i}{2064} \qquad (2)$$

where

P_i = probability of selection of the ith PSU.

The segment selection probability in the ith PSU was thus $\frac{MOS_{1ij}}{IP_i}$ for high-minority segments and $\frac{MOS_{2ij}}{IP_i}$ for low-minority segments. It should be noted that the overall segment selection probability was independent of P_i.

2.2.2.2 Minimum segment size

The screening sampling rate within a segment was $11/MOS_{2ij}$ (in low-minority segments) and $33/MOS_{1ij}$ (in high-minority segments). Thus, in the low-minority stratum

$$H_{Oij} + H_{Mij} \geq 11. \qquad (3)$$

In the high-minority stratum

$$H_{Oij} + 3H_{Mij} \geq 33. \qquad (4)$$

or

$$H_{Oij}/3 + H_{Mij} \geq 11. \qquad (5)$$

The actual segment sizes had to be 11 households in low-minority areas, 11 households in high-minority areas with 33 percent minorities, and 22 households in segments with 25 percent minorities.

2.2.2.3 Segment sample selection

The first step in sampling segments was to extract block data from the PL94 file for the 101 PSUs in the sample. In the next step, blocks containing fewer than the minimum number of housing units required to select the third-stage sample were combined with other adjacent or nearby blocks to form the segments that served as second-stage sample units. Segments were sorted within each PSU according to the proportion of Black and Hispanic residents.[2] A systematic sample of segments was then selected with probability proportional to size. The systematic selection provided implicit stratification according to the proportion of minority residents in the segments. The sample of 2,064 segments included 869 high-minority and 1,195 low-minority segments.

2.2.2.4 TIGER maps

The National Adult Literacy Survey was one of the first sample surveys nationwide to use the Bureau of the Census's Topologically Integrated Geographical Encoding and Referencing (TIGER) System file for the production of segment maps. Segment maps are essential features of an area sample; they define and describe the sample segments, permitting field interviewers to locate the areas and list the housing units within the segments. In the past, segment maps were produced by hand, with clerks outlining the segments manually on maps purchased from the Census Bureau. This operation was slow, costly, and somewhat error-prone. The maps were of diverse sizes, resulting in problems of filing and storage. Street names were difficult to read on many of the maps.

The Census Bureau produced a system known as the TIGER file for the implementation of the 1990 census. The TIGER file digitized all intersections of geographic boundaries used in the 1990 census, including individual blocks. This information can be used to computer generate maps of selected blocks, combinations of blocks, or any other type of geography referred to in the census. Before the National Adult Literacy Survey began, Westat purchased a copy of the TIGER file and software to generate maps from the file and then developed additional software to facilitate its use for sample survey purposes. In the completely automated sampling process, sample blocks were selected from census summary tapes, and the block identifications were automatically fed into the TIGER file, which in turn generated the segment maps. This method of map production cost considerably less than the old method, was more accurate, and was much faster to implement. Because Westat developed much of the software, other useful features were included in the segment maps. For example, the maps were uniform in size, had sufficient detail to permit

---

[2] A serpentine sort executes multiple sorts within a stratum such that bordering sample units are the most similar with respect to the sort variables. This is accomplished by reversing the sort order within the segment groups.

street names to be read, had convenient map numbers automatically inserted, included small-scale maps of larger areas, showed segment locations within broader areas, and included certain data (based on the 1990 census) for quality control.

2.2.2.5 Listing sample segments

Westat field staff visited each sample segment and prepared a list of all housing units within the boundaries of the segment. (A total of 142 large segments from the national sample were subdivided before listing, with one part, or "chunk," selected at random for listing.) Table 2-4 provides the distribution of segments in the national sample, by segment size. As noted earlier in this section, segments consisted of census blocks or combinations of two or more adjacent blocks that could be accessed without crossing over census tract boundaries. Therefore, if the segments did not contain enough households to reach the minimum size established for that type of segment (see section 2.2.2.2), the measure of size was considered to be equal to the minimum measure of size.

Table 2-4. Distribution of segments in the national sample, by segment size*

Dwelling units	Frequency	Percent	Cumulative frequency	Cumulative percent
0-19	8	0.4	8	0.4
20-29	1	0.0	9	0.4
30-39	12	0.6	21	1.0
40-49	35	1.7	56	2.7
50-59	100	4.8	156	7.6
60-69	282	13.7	438	21.2
70-79	264	12.8	702	34.0
80-89	196	9.5	898	43.5
90-99	129	6.2	1,027	49.8
100-119	211	10.2	1,238	60.0
120-149	208	10.1	1,446	70.1
150-199	186	9.0	1,632	79.1
200-249	102	4.9	1,734	84.0
250-299	103	5.0	1,837	89.0
300-399	168	8.1	2,005	97.1
400-499	51	2.5	2,056	99.6
500-699	7	0.3	2,063	100.0
700-799	1	0.0	2,064	100.0

* The frequencies reported in this table are the actual numbers of dwelling units listed in the selected segments. Large segments were subdivided and one section was selected at random for listing.

2.2.3 Third-Stage Sample—Selecting Housing Units

The third stage of sampling for the national component involved sampling households within the selected segments. After selection, households were screened to determine whether they included any eligible respondents. In the low-minority segments, any household with at least one person age 16 or older was included in the sample. In the high-minority segments, all minority households with at least one person age 16 or older were retained in the sample, but only one-third of nonminority households (with at least one person age 16 or older) were included in the sample.

2.2.3.1 Within-segment sampling rate

The sampling rates within the low-minority segments were set to produce an average of 11 housing units per segment. In high-minority segments, the average was about 14 housing units. White, non-Hispanic households in high-minority segments were sub-sampled at a rate of about one-third, so that White, non-Hispanic adults from high-minority segments had the same overall sampling rate as those residing in low-minority segments. The within-segment sampling rate (i.e., the household sampling rate) in low-minority segments was

$$r_{2ij} = \frac{11}{MOS_{2ij}} \tag{6}$$

In high-minority segments, the sampling rate was

$$r_{1ij} = \frac{33}{MOS_{2ij}} \text{ for minority households} \tag{7}$$

$$r_{2ij} = \frac{11}{MOS_{1ij}} \text{ for other households} \tag{8}$$

If the number of housing units in the selected segments was the same in 1992 as in 1990, the number of selected households that remained in the sample for interview would be constant across all segments; that is, if in low-minority segments the number of households in segment ij was equal to $H_{Oij} + H_{Mij} = MOS_{2ij}$, the sample size was equal to

$$\frac{11}{MOS_{2ij}} \text{ x } (H_{Oij} + H_{Mij}) = 11. \tag{9}$$

In high-minority segments, the sample size was equal to

$$n_o = \frac{11}{MOS_{1ij}} \text{ x } H_{Oij} = \frac{11 H_{Oij}}{H_{Oij} + 3H_{Mij}} \tag{10}$$

$$n_M = \frac{33}{MOS_{1ij}} \text{ x } H_{Mij} = \frac{33 H_{Mij}}{H_{Oij} + 3 H_{Mij}}$$

$$n_o + n_M = \frac{11 (H_{Oij} + 3H_{Mij})}{H_{Oij} + 3 H_{Mij}} = 11$$

where

n_o is the number of non-minority households selected in a high-minority segment;
n_M is the number of minority households selected in a high minority segment.

The segment sizes would thus be constant, equal to 11.

However, segment sizes for the screening sample varied in the high-minority stratum. The screening sample in each segment was the rate at which minorities were selected. The sampling yield for the screening sample was thus MOS

$$\frac{33}{MOS_{1ij}} (H_{Oij} + H_{Mij}) = \frac{33}{H_{Oij} + 3H_{Mij}} (H_{Oij} + H_{Mij}) \tag{11}$$

$$= 11 \left[\frac{H_{Oij} + 3H_{Mij}}{H_{Oij} + 3H_{Mij}} + \frac{2H_{Oij}}{H_{Oij} + 3H_{Mij}} \right]$$

$$= 11 \left[1 + \frac{2H_{Oij}}{H_{Oij} + 3H_{Mij}} \right]$$

$$= 11 (1 + \frac{\frac{2 H_{Oij}}{MOS_{1ij}}}{\frac{H_{Oij}}{MOS} + \frac{3H_{mij}}{MOS_{1ij}}})$$

Since the cut-off point for the high-minority strata was 25 percent minorities, the proportion of minorities in a segment from a high-minority stratum ranged from 25 percent to 100 percent. Putting those values in the formula above gives a range for the screening sample of 11 to 22 households.

In the national sample, 24,522 households were selected. The following table provides the distribution of the selected households by census region.

Census region	Number of households
Northeast	4,676
Midwest	5,051
South	9,340
West	5,455
Total	24,522

2.2.3.2 Overall probabilities of selection

The overall probability of selection of households in low-minority segments was

$$P_{2} = P_{i} \frac{MOS_{2ij}}{I\,P_{i}} \frac{11}{MOS_{2ij}} = \frac{11}{I} \qquad (12)$$

In high-minority segments, the overall probability of selection for nonminority households was equal to

$$P_{2i} = P_{i} \frac{MOS_{1ij}}{I\,P_{i}} \frac{33}{MOS_{1ij}} \frac{1}{3} = \frac{11}{I} \qquad (13)$$

where *I* is the sampling interval.

For minority households in high-minority segments, the overall probability was

$$P_{1i} = P_{i} \frac{MOS_{1ij}}{I\,P_{i}} \frac{33}{MOS_{1ij}} = \frac{33}{I} \qquad (14)$$

2.2.3.3 Procedures for selecting missed structures and missed dwelling units

Entire structures may have been omitted from the initial segment listing, either because the lister made an error or because the structure was constructed in the interval between listing and interviewing. Additionally, listers may have missed dwelling units within a listed structure because they were instructed not to inquire about the number of units in most residential buildings in order to reduce listing costs. Instead, listers were told to list a structure that looked like a one-family residence as a one-family residence. However, a smaller number of buildings that looked like one-family residences may have been converted to multi-family residences. To compensate for this problem and identify missed households, Westat instructed interviewers to conduct two quality control procedures at the time of data collection. These procedures are described below.

Missed Structure Procedure. If the first dwelling unit on the completed listing sheet was selected for the sample, a segment recanvass to search for missed structures was conducted. If any missed structures were found, the dwelling units within each missed structure were selected if the number of units within the structure was less than or equal to 10. If the number of units was greater than 10, 10 dwelling units were selected at random.

Missed Dwelling Unit Procedure. If the first (or only) dwelling unit on the completed listing sheet was selected for the sample, the interviewer inquired at the sample unit about any additional units in the building. If any missed dwelling units were found, then all missed units were selected if the number of missed dwelling units within the structure was less than or equal to 10. If the number of missed dwelling units within the structure was greater than 10, all missed units were listed and a sample was selected from the listing.

The increase in the total number of assessments and the effects of differential weights were considered when determining the probabilities with which to select these dwelling units. The overall goal was to control the increase in the total number of assessments within a segment so that no more than (approximately) double the number of persons originally expected were selected in a segment.

2.2.4 Fourth-Stage Sample—Selecting Persons Age 16 or Older

A list of household members was obtained during the screener interview conducted at each sample household. Interviewers listed the household members in descending order of age. A computer-generated sampling message attached in advance to each questionnaire contained instructions on which household members to choose for an interview. The following table illustrates a typical sampling message:

Number of eligible persons in household	Choose the following person for interview
1	First
2	Second
3	Second
4	First and third
5	First and fifth
6	Third and sixth
Etc.	

Because the sampling messages varied from household to household, each household member had the same chance of selection within each size of household group. One adult was sampled randomly from households with fewer than four eligible persons. In households with four or more eligible persons, two adults were selected. The selection of two adults in households with four or more eligible persons prevented a substantial increase in variances due to high weights resulting from the selection of one person in households with large numbers of eligible persons.

Because non-Black, non-Hispanic persons were undersampled in segments designated as high minority, each individual was classified into a race/ethnicity class during the screening interview so that the subsampling procedure for non-Black, non-Hispanic persons could be implemented. Because most U.S. households contain persons of the same race/ethnicity group, a race/ethnicity category was also assigned to each household and the subsampling procedure was carried out based on the race/ethnicity of the household. The household classification was based on the race/ethnicity of the person designated as the head of household, defined as the person who owns or rents the dwelling unit. If the screener respondent could not identify a head of household, the race/ethnicity of the first person listed on the household roster was used as the race/ethnicity of the household. This procedure made the sample screening and selection less complicated and reduced the chance of sample selection errors during the data collection.

The subsampling of nonminority households in high-minority segments was carried out using a sampling message that was attached to the questionnaires for a randomly selected two-thirds of the households in high-minority segments.

2.3 THE NON-INCENTIVE SAMPLE

At the request of the Office of Management and Budget, a subsample of segments was selected to produce about 1,000 completed interviews with respondents who were not offered the $20 incentive. A field test experiment carried out before the main survey showed lower response rates for the non-incentive group than for those who received incentives. The lower response rates were taken into account when selecting the segment sample for the non-incentive experiment.

A subsample of 155 segments was selected randomly from the 2,064 segments in the national sample, including 65 high- and 90 low-minority segments. This subsample contained 1,812 households and was expected to yield approximately 1,000 completed interviews with respondents who received no incentives.

The role of incentives is discussed in more detail in Chapter 10 of this report.

2.4 SAMPLING FOR THE STATE LITERACY SURVEYS

The National Adult Literacy Survey provided an opportunity for state officials to request that supplementary adult literacy surveys be conducted within their states, to provide state-level estimates of adult literacy skills that are reliable, valid, and comparable to national estimates. A sample of about 1,000 interviewed persons was used to supplement the national sample in each of the 11 states participating in the program. This sample size was estimated to be sufficient to provide adequate precision for most anticipated analyses.

Participants in the state component were selected through a process nearly identical to that used for the national component, where the units at each stage of sample selection represented a particular state rather than the entire United States. The two principal differences between the sample designs for the national and the state surveys were that (1) Black and Hispanic adults were not oversampled in the state surveys and (2) the respondent universe consisted of adults ages 16–64 (vs. adults age 16 and older for the national survey).

2.4.1 Sample of PSUs

The first-stage primary sampling units, or PSUs, for a state consisted of geographic clusters of one or more adjacent counties within the state. With a few exceptions, the PSUs were identical to those used in the national sample. The exceptions were the national PSUs that crossed state boundaries, which were subdivided for the state sample. Each PSU was assigned to a stratum (i.e., groups of PSUs with similar characteristics) and one PSU was selected within each stratum. The following characteristics were used to stratify the PSUs: whether the PSU was within an MSA as defined for the 1990 census; the percentage of the population in the PSU who were Black and/or Hispanic; and the population size of the PSU. Per capita income was also used wherever possible. In some states, the number of strata that could be created precluded the effective use of all four stratifying characteristics.

One PSU was selected from each stratum with a probability proportional to the PSU's 1990 population. The number of sample PSUs per state varied from 8 to 12, with smaller numbers of PSUs in states with one or more very large PSUs that were chosen with certainty.

2.4.2 Sample of Segments

The second-stage sampling units consisted of census blocks or groups of blocks within the selected PSUs. Adjacent blocks were combined whenever necessary to ensure that each segment had a minimum of 20 housing units per segment. In each state, 167 segments were selected across the PSUs. The selection was systematic and with probability proportional to size, where the measure of size was the number of year-round housing units within the segment.

The sampling interval for the selection of segments, I, was computed as

$$I = \frac{\sum_{I_s} MOS_{2ij}/P_i}{167} \qquad (15)$$

where

MOS_{2ij} = the measure of size for the j^{th} segment in the i^{th} PSU (note that this is equivalent to the low-minority segment measure of size in the national component) and

P_i = probability of selection of the i^{th} PSU.

The PL94 data tapes from the 1990 census were used to define the segments within each PSU. Segments were stratified according to the percentage of minority (Black and Hispanic) residents before selection.

2.4.3 Sample of Housing Units

The third stage of sampling involved the selection of households within segments. Westat field staff visited the 167 selected segments and prepared a list of all housing units within the boundaries of each segment. Segment boundaries were determined by the 1990 census block maps (i.e., the TIGER maps). The segment listings were sent to Westat, where a sample of about 11 housing units was selected per segment. Interviewers visited these housing units, determined which were occupied, and obtained a roster of household members. The same quality control procedures as in the national sample were used to compensate for missed structures and missed dwelling units within listed structures.

2.4.4 Sample of Persons

One or two adults ages 16–64 were selected from the list of household members obtained during the household screening. The selection procedure was similar to the one used in the national sample. One person was selected at random from households with fewer than four eligible members; two persons were selected from households with four or more eligible persons. Interviewers listed the eligible household members in descending age order. The interviewers then identified the one or two household members for interview based on computer-generated sampling messages that had been attached to each questionnaire in advance.

2.5 WEIGHTED AND UNWEIGHTED RESPONSE RATES

Unweighted response rates are indicators of how well the survey operations were carried out. They are useful during the survey as part of the quality control process and at the completion of field work as a measure of success. However, weighted response rates are more appropriate in examining the potential effect of nonresponse on statistics. Because the literacy estimates are based on weighted data, weighted response rates are better clues to potential data quality problems. Table 2-5 provides the weighted and unweighted response rates for the survey. Note that for the National Adult Literacy Survey the weighted and unweighted response rates are almost identical. Chapter 3 includes a detailed discussion of the weighting procedures used in the National Adult Literacy Survey.

Table 2-5. Screener, background questionnaire, and exercise booklet response rates for the National Adult Literacy Survey, by respondent characteristics for all sample types

Survey component and subgroup	Unweighted (%)	Weighted (%)*
Screener	89.1	--
Background questionnaire		
All respondents	81.0	80.5
Age		
16-24	85.0	85.5
25-44	82.8	82.3
45-64	78.7	78.1
65+	77.4	74.9
Sex		
Male	77.9	77.9
Female	83.5	82.7
Race/ethnicity		
Hispanic	81.7	82.3
Black, non-Hispanic	84.6	84.0
White and other	80.2	79.9
Exercise booklet		
All respondents	95.9	95.9
Age		
16-24	98.2	98.6
25-44	96.7	96.7
45-64	94.6	94.5
65+	89.0	
Sex		
Male	95.7	95.6
Female	96.0	96.2
Race/ethnicity		
Hispanic	95.0	95.4
Black, non-Hispanic	94.3	94.8
White and other	96.3	96.1
Education level		
Some or no high school	94.0	93.9
High school graduate/GED**	95.4	95.3
Some college or vocational education	96.7	97.0
College graduate or advanced degree	97.1	97.0

* The weighted response rates were calculated by applying the sampling weight to each individual to account for his/her probability of selection into the sample. Weighted response rates were computed only for screened households (the probability of selection is not known for persons in households that were not screened).

** GED = General Educational Development certificate

2.6 SAMPLING FOR THE PRISON SURVEY

For the survey of the prison population, background interviews were completed with 1,147 persons. The survey used a two-stage sample design. The first-stage unit, or PSU, was a state or Federal adult correctional facility selected with probability proportional to size, where the measure of size was the size of the inmate population. The second-stage unit was an inmate within a sample facility. Inmates were selected with a probability inversely proportional to the facility's population size, so that the product of the first- and second-stage selection probabilities would be constant. The selection rates were designed to produce an average of about 12 assessments per facility. In practice, this number varied because of differences between the anticipated and actual sizes of the inmate populations.

Although the sample design was intended to provide a constant overall probability of selection across all inmates, inmate selection probabilities were lowered in a few facilities because of operational constraints. In facilities with high rates of population growth, the sample size to yield a constant selection probability exceeded the maximum allowable number of interviews (22). Because the sample sizes in these facilities had to be truncated to 22, the overall selection probabilities were lower. Sections 2.6.1 and 2.6.2 describe the procedures for selecting correctional facilities and inmates, respectively.

2.6.1 Sample of Correctional Facilities

In the first stage of sampling, a sample of Federal and state adult correctional facilities was selected. The correctional units in multi-location facilities were sub-sampled, and one correctional unit was selected from each multi-location facility. It was estimated that, with a sample of approximately 15 inmates from each facility, a maximum of 96 facilities would be necessary to produce the required number of completed background interviews (1,000). This estimate was based on the assumptions that approximately 80 of 96 facilities (83 percent) would cooperate and that, on average, interviews would be completed with approximately 12 to 13 inmates in each of the cooperating facilities. However, early successes in gaining the cooperation of selected facilities indicated that response rates much higher than the anticipated 83 percent were likely. Therefore, a random subsample of eight facilities was deselected and set aside as a reserve sample. Of the 88 facilities selected for data collection, 87 (one of which was discovered to be two facilities) agreed to cooperate, and one facility was determined to be ineligible. The gain of one facility offset the loss of one facility due to ineligibility, making the number of eligible facilities 88. Therefore, it was not necessary to use the reserve sample.

2.6.1.1 Sampling frame and selection of correctional facilities

The sampling frame for the correctional facilities was based on the 1990 census of Federal and state prisons. The data in the frame were updated to mid-1991. State adult correctional and Federal adult correctional facilities were extracted from the census file.[3]

---

[3] The youth offender facilities is a category under the state adult prisons.

The sample of correctional facilities was drawn from the correctional facilities frame. The facilities in the frame were stratified on the basis of their characteristics using implicit stratification. That is, the facilities were placed in a sort order according to these characteristics and were selected systematically. The following variables were used in the sort:

1) State or Federal;
2) Region: Northeast, Midwest, South, West;
3) Sex of inmates: male only, both sexes, female only; and
4) Type of facility:
    a) For state facilities, the categories in the sort order were maximum and closed security; medium security; minimum security; classification, diagnostic, and reception center; medical facility and hospital; work-release/prerelease; and youthful offender facility.
    b) For Federal facilities, the categories in the sort order were U.S. penitentiary, Federal correctional institution, federal prison camp, metropolitan correctional center, federal detention center, metropolitan detention center, federal medical center, community correctional center, and other.

The facilities were sorted first according to whether they were federal or state facilities; then by region, inmate gender composition within region, and type of facility within inmate gender composition; and, finally, by the size of the facility's inmate population within type of facility. A serpentine sort order was used for the last three variables. That is, the direction of the sort for inmate gender composition alternated between region categories, and the direction of the sort for type of facility alternated between inmate gender composition categories.

From this sorted list, the sample of facilities was drawn by taking a systematic sample with probabilities proportional to the number of inmates in the facility. The number of inmates in a facility was taken as its measure of size. The reserve sample of eight facilities was drawn by taking a systematic sample, with equal probabilities of selection, from the 96 sample facilities.

Table 2-6 shows the numbers of correctional facilities in the sample (excluding the reserve units), as well as facilities and inmates in the sampling frame, by stratification variables.

Table 2-6. Number of facilities and inmates included in the survey of the prison population, by stratification variables*

	Sample		Sampling frame			
	Facilities		Facilities		Inmates	
Stratification variable	Number	%	Number	%	Number	%
Facility Type						
Total	88	100.0	1,345	100.0	712,141	100.0
State	81	92.0	1,250	92.9	654,646	91.9
Federal	7	8.0	95	7.1	57,495	8.1
State Facilities						
Total	81	100.0	1,250	100.0	654,646	100.0
Region						
Northeast	14	17.3	195	15.6	117,221	17.9
Midwest	18	22.2	264	21.1	141,988	21.7
South	30	37.0	546	43.7	249,705	38.1
West	19	23.5	245	19.6	145,732	22.3
Facility type						
Maximum security	24	29.6	186	14.9	197,230	30.1
Medium security	37	45.7	392	31.4	298,380	45.6
Minimum security	10	12.3	334	26.7	83,909	12.8
Classification, Diagnostic, and Reception center	4	4.9	43	3.4	32,896	5.0
Medical facility	1	1.2	3	0.2	7,653	1.2
Work-release Prelease center	3	3.7	265	21.2	20,505	3.1
Youthful offender Facility	2	2.5	27	2.2	14,073	2.1
Sex of inmates						
Male only	73	90.1	1,027	82.2	584,539	89.3
Both sexes	5	6.2	117	9.4	43,183	6.6
Female only	3	3.7	106	8.5	26,924	4.1
Federal facilities						
Total	7	100.0	95	100.0	57,495	100.0
Region						
Northeast	1	14.3	13	13.7	8,339	14.5
Midwest	1	14.3	15	15.8	10,913	19.0
South	3	42.8	50	52.6	27,964	48.6
West	2	28.6	17	17.9	10,279	17.9

Table 2-6. Number of facilities and inmates included in the survey of the prison population, by stratification variables* – continued

	Sample		Sampling frame			
	Facilities		Facilities		Inmates	
Stratification variable	Number	%	Number	%	Number	%
Facility type						
U.S. penitentiary	1	14.3	6	6.3	7,360	12.8
Federal correctional Institution	3	42.9	32	33.7	29,865	51.9
Federal prison camp	3	42.8	34	35.8	11,373	19.8
Metropolitan correctional center	0	0.0	4	4.2	3,400	5.9
Federal detention center	0	0.0	6	6.3	1,648	2.9
Metropolitan detention center	0	0.0	1	1.0	867	1.5
Federal medical center	0	0.0	2	2.1	1,679	2.9
Community correctional center	0	0.0	7	7.4	787	1.4
Other	0	0.0	3	3.2	516	0.9
Sex of Inmates						
Male	6	85.7	72	75.8	47,281	82.2
Both sexes	1	14.3	19	20.0	8,808	15.3
Female	0	0.0	4	4.2	1,406	2.5

*Excludes reserve sample.

2.6.2 Selection of Inmates Within Facilities

An upper bound of 22 inmates per facility was used to determine the inmate sample sizes for the correctional facilities. This upper bound was dictated by the practical limits on interviewing a large number of inmates per facility. First, the expected inmate sample sizes for cooperating facilities were computed under a self-weighting design to yield a total of 1,500 inmates. If a facility's expected sample size exceeded 22, it was truncated to 22, and the sample sizes for the other facilities were inflated to yield a total expected inmate sample of 1,500. This iterative process continued until there was no facility with an expected inmate sample size greater than 22, and the expected inmate sample sizes summed to 1,500 over all cooperating facilities.

Because of the uncertainty concerning inmate response rates and their availability for interview, the sample of facilities was randomly divided into two waves. The first wave included 30 percent of the facilities. The outcomes of wave 1 (in terms of response rates and inmate availability) were used to set the sampling rates for wave 2.

The selection of inmates was conducted within each facility using a list of names obtained from facility administrators. The interviewers received forms to complete and instructions that they were required to follow when sampling inmates from the lists.

Chapter 3

WEIGHTING AND POPULATION ESTIMATES

Leyla Mohadjer, John Burke, James Green, and Joseph Waksberg, Westat, Inc.

3.1 GOALS OF WEIGHTING

Sample weights were produced for the National Adult Literacy Survey respondents who completed the exercise booklet; those who could not start the exercises because of a language barrier, a physical or mental disability, or a reading or writing barrier; and those who refused to complete the exercises but had completed background questionnaires. Separate sets of weights were computed for the incentive and non-incentive samples (refer to section 2.3 and Chapter 10 for a description of the non-incentive sample).

The purpose of calculating sample weights for the National Adult Literacy Survey was to permit inferences from persons included in the sample to the populations from which they were drawn, and to have the tabulations reflect estimates of the population totals. Sample weighting was carried out to accomplish the following five objectives:

1) To permit unbiased estimates, taking account of the fact that all persons in the population did not have the same probability of selection;
2) To combine the state and national samples in an efficient manner;
3) To bring data up to the dimensions of the population totals;
4) To use auxiliary data on known population characteristics in such a way as to reduce sampling errors; and
5) To minimize biases arising from differences between cooperating and non-cooperating persons in the sample.

Objective 1 was accomplished by computing base weights for the persons selected into the sample. To produce unbiased estimates, different weights must be used for various subsets of the population, whenever these subsets have been sampled at different rates. Weighting was required to account for the oversampling of black and Hispanic persons in high-minority segments of the national sample. Furthermore, the survey specifications called for the selection of one person in households with fewer than four eligible members and two persons in households with four or more eligible members. Using this approach, members of households with only one eligible member had twice the chance of selection of those in households with two eligible members, three times the chance of selection of those in households with three eligible members, etc. Weighting was needed in these situations to prevent potentially serious biases.

The base weight was calculated as the reciprocal of a respondent's final probability of selection. For the household sample, it was computed as the product of the inverse of probabilities of selection at the primary sampling unit (PSU), segment, household, and person levels. For the prison sample, the base weight was equal to the reciprocal of the product of the selection probabilities for the facility and the inmate within the facility. Section 3.2.2 provides a summary of the base weight computation.

The second objective of weighting was to provide composite weights for the respondents in the 11 state samples and the respondents in the national sample PSUs in the 11 states. The national and state components applied the same sampling procedures in terms of stratification method, PSU construction, sample design, and selection at the various stages of sampling. Furthermore, the same forms were used to screen households and to collect background information and literacy assessment data in the state and national surveys. To take full advantage of this comparability, the samples were combined to produce both state- and national-level statistics. The advantage of compositing the samples was the increased sample size, which improved the precision of both state and national estimates. It should be noted that composite estimates apply only to persons ages 16–64, because data for persons age 65 and older came only from the national sample. Section 3.2.4 describes the composite estimation procedures used for the National Adult Literacy Survey.

For the household components, the post-stratified base weight was multiplied by a compositing factor that combined the national and state component data in an optimal manner, considering the differences in sample size and sampling error between the two components. Up to four different compositing factors were used in each of the 11 participating states, and a pseudo factor (equal to 1) was used for all persons age 65 and older and for national component records from outside of the 11 states. The product of the post-stratified base weight and the compositing factor for a record was the composite weight. A particular state analysis can include data from all respondents, age 16 and older, in that state. However, the sampling error for state estimates will increase with the inclusion of records for respondents over age 64, because these records came from the national component only.

Objectives 3, 4, and 5 were accomplished in one step by adjusting for nonresponse through post-stratification and raking[1] to adjusted 1990 census totals. If every selected household had agreed to complete the screener, and every selected person had agreed to complete the background questionnaire and the exercise booklet, weighted estimates based on the data would be approximately unbiased (from a sampling point of view). However, nonresponse occurs in any survey operation, even when participation is not voluntary. The best approach to minimizing nonresponse bias is to plan and implement field procedures that maintain high cooperation rates. For example, the payment of a $20 incentive in the household survey and repeated callbacks for refusal conversion were very effective in reducing nonresponse, and thus nonresponse bias. However, because some nonresponse occurs even with the best strategies, adjustments are always necessary to avoid potential nonresponse bias.

---

[1] Raking is a special kind of poststratification in which the weights of the adjustment cells are adjusted in such a way that the weighted sample marginal totals correspond to known population totals.

Although the data collection was carried out in 1992, adjusted 1990 census data (adjusted for undercoverage) were used for poststratification. Undercount rates estimated by the U.S. Bureau of the Census were applied to the 1990 census count to correct for the undercoverage of some population subgroups. It was concluded that the estimates would not have been improved by extrapolating 1990 census data to the 1992 estimates of the population.

The composite weights were raked so that numerous totals calculated with the resulting full sample weights would agree with the 1990 census totals, adjusted for undercount. The cells used for the raking were defined to the finest combination of age, education level, and race/ethnicity that the data would allow. Raking adjustment factors were calculated separately for each of the 11 states and then for the remainder of the United States. Section 3.2.5 describes the details of the poststratification and raking approaches. Demographic variables that were critical to the weighting were re-coded and imputed, if necessary, before the calculation of base weights.

Full-sample and replicate weights were calculated for each record to facilitate the computation of unbiased estimates and their standard errors. The full-sample and replicate weights for the household components were calculated as the product of a record's post-stratified base weight and a compositing and raking factor.

The weighting procedures were repeated for 60 strategically constructed subsets from the records in the sample to create a set of replicate weights for variance estimation using the jackknife method. The replication scheme was designed to produce stable estimates of standard errors for the national and 11 individual state estimates.

The full-sample and replicate weights for the prison component were calculated as the product of a record's base weight and a nonresponse and raking factor. The base weight was calculated as the reciprocal of the final probability of selection for a respondent, which reflected the two stages of sampling (sampling facilities and sampling inmates within facilities). The base weights were then adjusted for nonresponse to reflect both facility and inmate nonresponse. The resulting nonresponse-adjusted weights were then raked to agree with independent estimates for certain subgroups of the population. The raking and nonresponse adjustment was performed independently by replicate.

3.2 CALCULATING SAMPLE WEIGHTS FOR THE HOUSEHOLD POPULATION

3.2.1 Preliminary Steps in Weighting

The data used in weighting underwent edit, frequency, and consistency checks to prevent any errors in the sample weights. The checks were performed on fields required for data weighting and were limited to records that required weights (i.e., records for respondents who completed the exercise booklet and those who failed to complete a screener.

The consistency checks also helped to identify any unusual values. Listings were prepared of records with missing values in any of the fields used in weighting. The listings showed the entire record: the respondent's identification number, age, date of birth (from the background questionnaire), sex, race/ethnicity, level of education, the race of the head of household, and the number of age-eligible members and respondents in the household. The printed listings were used to review the extent of missing data, identify the pattern of missing data, and prepare for imputation. The sex and race/ethnicity data from the screener and background questionnaire were also compared for consistency. Overall, these checks found little missing data and very few records with values that differed between the screener and the background questionnaire.

Most of the fields required for data weighting (race/ethnicity of the head of household; sex, age, race/ethnicity and education of the respondent) were at finer levels of detail than were necessary for the later steps of weighting. The data in these fields were, therefore, collapsed to the required levels. Most of these fields were present in both the screener and the background questionnaire, thereby providing two measures of the same item. The background questionnaire measure was preferred for all items except the race of the head of household, which was collected only on the screener. For the few cases in which the background questionnaire measure was missing, the screener measure was generally available and was used as a direct substitute. Frequencies were prepared for each item after collapsing and making direct substitutions to gauge the magnitude of the imputation task.

The amount of missing data remaining after substitution was small, making the imputation task fairly straightforward. The Westat imputation macro WESDECK was used to perform hot-deck imputation for particular combinations of fields that were missing. Imputation flags were created for each of the five critical fields to indicate whether the data were originally reported or were based on substitution or imputation via WESDECK. The imputed values were used only for the sample weighting process.

Several special cases required attention before the calculation of base weights. In some dwelling units, the number of eligible household members exceeded nine, the maximum allowable number on preprinted labels used by the interviewers for respondent selection. In these instances, field staff provided the total number of eligible household members to the main office, where statisticians randomly selected respondents for interview and relayed this information back to the field staff. Detailed records indicated the PSU, segment number, total number of eligible household members, and number of respondents selected in each dwelling unit. This information was retrieved and attached to each of these records before the calculation of base weights.

Some additional dwelling units came into the sample as part of the missed structure and missed dwelling unit procedures (refer to section 2.2.3.3 for more information), which allow units that were missed in the segment listing activities to be included in the sample with a known probability of selection. All missed dwelling units within a segment were included unless the total number of missed units in the segment was unusually large, in which case a sample of missed dwelling units was taken. Detailed records indicated the PSU, segment, number of missed dwelling units selected, and total number of missed dwelling units whenever a sample of missed units was selected. This information was retrieved and attached to each of these records prior to the calculation of base weights.

A few final checks were run before base weight calculation to ensure the availability and validity of all fields required by the base weights program (fields created for the special cases mentioned above and fields for the total number of age-eligible household members and the number of sample persons for each dwelling unit). A detailed description of base weight computation is provided in the next section.

3.2.2 Computing Base Weights

A base weight was calculated for each record. The base weight was initially computed as the reciprocal of the product of the probabilities of selection at each stage of sampling (as given in section 2.2.3.2). The base weight reflected the probabilities of selection at the PSU, segment, dwelling unit, and respondent levels. The final base weight included adjustments to reflect the selection of the reserve sample (see section 2.2.2), the selection of missed dwelling units (see section 2.2.3.3), and the chunking process conducted during the listing of the segments (section 2.2.2.5), and to account for the subsample of segments assigned to the non-incentive experiment (section 2.3) and the sub-sampling of respondents within households (section 2.2.4). The base weight was given by

$$W_{hij} = \frac{1}{P_{ij}} R\, k\, h_i\, C_i\, S_j \qquad (1)$$

where

P_{ij} = the initial probability of selection of household j in segment i;

R = the adjustment factor for the selection of the reserve sample;

k = the adjustment factor to reflect the sub-sampling of the non-incentive sample;

h_i = the adjustment factor for the addition of missed structures and dwelling units in segment i;

C_i = the adjustment factor to reflect the chunking of the segments during the listing operation; and

S_j = the factor to reflect the sub-sampling of persons in household j with multiple eligible members.

Twelve respondents in the national sample had extremely high base weights resulting from various features of the design. The base weights of these respondents were trimmed down to about three times the mean value of the base weights to avoid unnecessary increases in variances of estimates from the National Adult Literacy Survey.

3.2.3 Nonresponse Adjustments and Poststratification

Before compositing the national and state samples, the base weights for each sample were post-stratified separately to known population totals. This first-level poststratification provided sampling weights with lower variation and adjusted for nonresponse. Poststratification implicitly adjusts for unit nonresponse through adjustments to the weights of the responding units. Typically, the adjustments are made for subgroups of the sample that are likely to be quite different or for subgroups with high nonresponse rates. Poststratification is appropriate when population totals are known for the subgroups, or weighting classes, of the sample.

For purposes of poststratification, the entire sample was partitioned into classes, with the classification based on available survey data from respondents. Each class contained sample persons with the survey characteristics provided below. The adjustment was then implemented within each weighting class. The national and state records were split into 45 mutually exclusive and exhaustive groups, according to the state the record came from, whether the record came from the national or a state sample, and whether the record came from a PSU that was included in the national sample with certainty. The 45 groups were defined as follows:

Groups 1–11	State records from PSUs that were not selected with certainty for the national component, separated by state;
Groups 12–22	State records from PSUs that were selected with certainty for the national component, separated by state;
Groups 23–33	National records from one of the states participating in the state survey, from PSUs that were not selected with certainty for the national component, separated by state;
Groups 34–44	National records from one of the states participating in the state survey, from PSUs that were selected with certainty for the national component, separated by state; and
Group 45	National records from states not participating in the state survey.

State records were post-stratified separately from national records to provide a common base for applying the composite weighting factors. Population totals were calculated separately for each distinct group, based on 1990 census figures adjusted for undercount, thereby providing the control totals for poststratification. (More detail on poststratification totals is presented in section 3.2.5.)

A post-stratified base weight was calculated for each person in the sample as follows:

$$WPS_{hi} = Wb_{hi} \left[\frac{NT_h}{\sum_{i=1}^{n_h} Wb_{hi}} \right] \tag{2}$$

where

WPS_{hi} = the post-stratified base weight for the ith person record in the hth group;

Wb_{hi} = the base weight for the ith person record in the hth group;

NT_h = the population total for the hth group; and

n_h = the number of respondents in the hth group.

3.2.4 Compositing Data from the National and State Components

3.2.4.1 Composite estimation procedure

Composite estimates were developed so that the National Adult Literacy Survey data could be used to produce both state and national statistics. The original plan was to consider the national and state samples as two separate surveys, so that national statistics would be prepared from the national sample only and state data would be prepared from the state samples only. Upon reconsideration, it was clear that sampling error would be reduced by combining the state and national samples for each state that participated in the state survey. The combined sample had the advantages of producing a single database for state and national statistics and improving precision.

The method of combining data from the state and national samples is referred to as composite estimation. The composite estimation procedure and issues associated with the choice of composite weights for the national and state samples are discussed in the following sections.

The composite estimator for the national/state sample is given by

$$\hat{Y}_i = \beta_i \hat{Y}_{st} + (1 - \beta_i) \hat{Y}_{nt} \tag{3}$$

where

$\hat{Y}_i$ = the composite estimate for variable Y in state i;

β_i = the composite factor for state i ($0 < \beta_i < 1$);

$\hat{Y}_{st}$ = the estimate of Y coming from the state sample; and

$\hat{Y}_{nt}$ = the estimate of Y coming from the national sample.

The variance of a composite estimator will be smaller than the variance of both the national and state estimates if appropriate composite factors are used. Optimal factors can be found when unbiased estimators exist for the two components and approximate estimates of their variances are available. It should be noted that a composite estimator will produce unbiased estimates for any value of β_i. The optimum value of β_i is the one that results in the lowest variance. However, there is generally only a slight loss in efficiency if a reasonable approximation of the optimum value of β_i is used. In most practical situations (including the national and state components of the National Adult Literacy Survey), approximations are necessary because there is insufficient information available to provide the optimal value of β_i when sample weights are produced.

As stated earlier, the national and state samples were selected independently, and each could thus produce unbiased estimates of sub-domain statistics for persons 16–64 years of age. Therefore, factors could be derived to produce composite estimators with variances that were smaller than those of either of the two estimates. For statistic Y, the optimal composite factor for state i is

$$\beta_i = \frac{V(\hat{Y}_{nt})}{V(\hat{Y}_{nt}) + V(\hat{Y}_{st})} \qquad (4)$$

where

$V(\hat{Y}_{nt})$ = the variance of the estimate of Y coming from the national sample; and

$V(\hat{Y}_{st})$ = the variance of the estimate of Y coming from the state sample.

A different optimal value of β_i might be found for each statistic of interest. However, data analyses would be complicated if item-specific values of β_i were used, because items would not add up to totals, or totals derived by summing different items would not agree. Consequently, the goal for the National Adult Literacy Survey was to associate with each person in the sample a single compositing factor that, while not precisely optimal for any particular statistic, would be robust enough to enhance the precision of virtually all composited statistics. This objective was accomplished by focusing on aspects of the sample design that were likely to affect the variance, regardless of the choice of statistic. Under simple random sampling, the variance of the estimator is inversely proportional to the sample size, and the expression for β_i simplifies to the following:

$$\beta_i = \frac{n_{st}}{n_{st} + n_{nt}} \tag{5}$$

where

n_{st} = the number of respondents age 16-64 in the state sample; and

n_{nt} = the number of respondents age 16-64 in the national sample.

Because of the complexity of the National Adult Literacy Survey sample design, it was useful to think of deriving β_i in terms of the effective sample size, i.e., the actual sample size divided by the design effect. Three aspects of the survey design tended to inflate the design effect and thereby reduce the effective sample size: clustering, stratification, and the differential sampling rates used for black and Hispanic adults.

In both the national and state components, clustering occurred at the PSU and segment levels and, to a trivial extent, at the household level, where two respondents were sampled in a small proportion of households. Geographic clustering kept the cost of survey administration down but reduced the effective sample size because of within-PSU and within-segment intraclass correlations. For example, in the Current Population Survey, which has a PSU and segment sample design similar to that of the National Adult Literacy Survey, the within-PSU and within-segment intraclass correlations have been estimated to average about 0.00075 and 0.042, respectively (Train et al., 1978). It seemed reasonable to use these values as approximations of intraclass correlations for the national and state components of the National Adult Literacy Survey.

Ordinarily, stratification enhances sample efficiency, but the national PSU sample was designed to optimize the precision of national estimates. As a result, stratum boundaries did not always conform with state boundaries; in fact, because PSUs sometimes contained counties from more than a single state, the measure of size used for PSU sample selection was not always optimal for producing state estimates. This aspect of the national design affected the variances of the state-level estimates coming from the non-certainty PSUs included in the national sample. (Note that stratum boundaries do not cause any problem for PSUs selected with certainty, because they are self-representing.)

In the national sample, minority households were oversampled in segments containing a high proportion of black and Hispanic households. This practice introduced variability in the weights and increased the design effect. Minority households were not oversampled in the state survey. A separate source of variability in weights for both the national and state samples was the within-household sampling of persons, although this variability was dampened somewhat by increasing the sample size to two persons in households containing four or more eligible adults.

To best reflect the influence of these design aspects on the effective sample size, distinct compositing factors were derived for up to four subsets of data in each participating state. Those subsets were defined according to (1) whether or not the data came from a PSU chosen with certainty for the national sample and (2) whether or not the respondent was black or Hispanic.

3.2.4.2 Deriving the PSU design effect

As mentioned in the previous section, the national PSU sample was not designed to maximize the efficiency of state-level estimates. To estimate the relative loss of efficiency for state data resulting from the inclusion of the national non-certainty PSUs, special tabulations were produced for each of the 11 participating states. The analysis was based on a variable that was likely to be correlated with literacy at the PSU level: the percentage of persons age 25 or older who had 0–8 years of schooling. Although the use of 1990 census data would have been preferable, only 1980 figures were available at the time.

First, all possible PSU samples under the national sample design were enumerated, and the between-PSU variances were computed for the estimated percentage using a Taylor series approximation. This process was repeated for the state design. These variances, which are presented in the third column of Table 3-1, were used to calculate provisional compositing factors that would have been appropriate had no within-PSU sampling been performed. These compositing factors reflect the limitations of the national stratification procedures for producing efficient state estimates. The table shows that the national design was quite adequate for producing state estimates in California but was greatly deficient in Louisiana.

Under the hypothesis that the national and state designs were equally efficient, another set of compositing factors, based strictly on the counts of PSUs (excluding the certainty PSUs in the national sample), was computed. These figures are presented in the fifth and sixth columns of Table 3-1. A factor similar to a design effect was computed by taking the quotient of the ratio of the state and national compositing factors derived using the two approaches:

$$F_{ij} = \left[\frac{\beta_{\text{Between PSU variance}}}{(1 - \beta_{\text{Between PSU variance}})} \right] \Big/ \left[\frac{\beta_{\text{PSU count}}}{(1 - \beta_{\text{PSU count}})} \right] \tag{6}$$

$$= \left[\frac{\text{National between-PSU variance}}{\text{number of national PSUs}} \right] \Big/ \left[\frac{\text{State between-PSU variance}}{\text{number of state PSUs}} \right]$$

This factor plays a role in calculating the effective sample size, as described in the next section.

3.2.4.3 Estimating composite factors

For data collected in PSUs selected with certainty for both the national and state samples, the effective sample size was estimated as:

$$n_{eff}^{ijk} = \frac{n_{ijk}}{1 + (\bar{n}_{ijk} - 1)\rho_1 + V^2_{w_{ijk}}} \qquad (7)$$

where

i = a participating state;
j = national or state sample;
k = minority (black or Hispanic) or non-minority;
n_{ijk} = total number of respondents ages 16-64
$\bar{n}_{ijk}$ = mean number of respondents per segment;
ρ_1 = 0.042, the intraclass correlation within segment, assumed to be equal to the Current Population Survey average and to be constant across states; and
$V^2_{w_{ijk}}$ = the relvariance[2] of the weights.

---

[2] Relvariance, short for relative variance, is calculated by dividing the variance on an estimate by the squared value of the estimate.

Table 3-1. Between-PSU variance and provisional compositing factors for the National Adult Literacy Survey national and state PSU sample designs

State	Data source	Between variance*	Provisional compositing factors	PSU count†	Provisional compositing factors	F_{ij}**
California	National	0.000498	0.4644	4	0.5000	1.15
	State	0.000432	0.5356	4	0.5000	1.00
Illinois	National	0.001375	0.1735	3	0.3750	2.86
	State	0.000289	0.8265	5	0.6250	1.00
Indiana	National	0.000401	0.0865	4	0.2500	3.52
	State	0.000038	0.9135	12	0.7500	1.00
Iowa	National	0.001812	0.0324	2	0.1429	4.97
	State	0.000061	0.9676	12	0.8571	1.00
Louisiana	National	0.002499	0.0210	1	0.1000	5.19
	State	0.000053	0.9790	9	0.9000	1.00
New Jersey	National	0.000430	0.0000	4	0.2857	1.00
	State	0.000000	1.0000	10	0.7143	1.00
New York	National	0.000127	0.3964	2	0.3333	0.76
	State	0.000083	0.6037	4	0.6667	1.00
Ohio	National	0.000140	0.1703	5	0.2941	2.03
	State	0.000029	0.8297	12	0.7059	1.00
Pennsylvania	National	0.000214	0.2571	4	0.3333	1.44
	State	0.000074	0.7429	8	0.6667	1.00
Texas	National	0.001482	0.1715	4	0.3333	1.44
	State	0.000307	0.8285	8	0.6667	1.00
Washington	National	0.000390	0.0681	1	0.1111	1.71
	State	0.000029	0.9319	8	0.8889	1.00

* Of the estimated percentage of persons 25 or older (1980) with 0-8 years of schooling.

† Excluding the National Adult Literacy Survey certainty PSUs.

** A design-effect-like factor descriptive of the relative inefficiency of the national PSU sample design for making state estimates

For data collected in other than the certainty PSUs included in the national sample, the effective sample size was estimated as

$$n_{eff}^{ijk} = \frac{n_{ijk}}{1 + (\bar{n}_{ijk} - 1)\rho_1 + (\bar{m}_{ijk} - 1)\rho_2 P_{ijk} F_{ij} + V_{w_{ijk}}^2} \tag{8}$$

where

i = a participating state;

j = national or state sample;

k = minority (black or Hispanic) or nonminority;

n_{ijk} = total number of respondents ages 16-64;

$\tilde{n}_{ijk}$ = mean number of respondents per segment ages 16-64;

ρ_1 = 0.042, the intraclass correlation within segment, assumed to be equal to the Current Population Survey average and to be constant across states

$\bar{m}_{ijk}$ = mean number of respondents per segment;

ρ_2 = 0.00075, the intraclass correlation within PSU, assumed to be equal to the CPS average and to be constant across states;

P_{ijk} = the proportion of respondents in non-certainty PSUs;

F_{ij} = a design-effect-like factor descriptive of the relative inefficiency of the national PSU sample design for making state estimates; and

$V^2_{w_{ijk}}$ = the relvariance of the weights.

Then an estimate of the optimal composite factor for state i is given by

$$\beta_{i(State)k} = \frac{n_{eff_{i(State)k}}}{n_{eff_{i(State)k}} + n_{eff_{i(National)k}}} \qquad (9)$$

$$\beta_{i(National)k} = 1 - \beta_{i(State)k} = \frac{n_{eff_{i(National)k}}}{n_{eff_{i(State)k}} + n_{eff_{i(National)k}}} \qquad (10)$$

Table 3-2 presents each of the quantities contained in the above formulas and the final compositing factors.

Table 3-2. Derivation of factors used to composite the National Adult Literacy Survey national and state data

State	National Certainty PSU	Race/ethnicity	Data source	Sample Size	Persons/ segment	Persons PSU	$P_{ijk\cdot}$	$F_{ij\cdot}$	Relvariance of weights	Effective design effect	Sample size	Compositing factor
California	No	black or Hispanic	National	196	3.5	49.0	1.0	1.2	0.3305	1.48	132.7	0.7098
			State	62	2.1	20.7	1.0	1.0	0.0804	1.14	54.2	0.2902
		Other	National	200	3.7	50.0	1.0	1.2	0.1393	1.30	154.4	0.4401
			State	260	4.7	65.0	1.0	1.0	0.1191	1.32	196.4	0.5599
	Yes	black or Hispanic	National	675	13.0	-	-	-	0.3666	1.87	361.0	0.6883
			State	226	7.5	-	-	-	0.1083	1.38	163.5	0.3117
		Other	National	414	8.3	-	-	-	0.1177	1.42	290.9	0.5232
			State	457	15.2	-	-	-	0.1261	1.72	265.1	0.4768
Illinois	No	black or Hispanic	National	56	4.3	18.7	1.0	2.9	0.3629	1.54	36.4	0.5968
			State	29	1.8	7.25	1.0	1.0	0.1414	1.18	24.6	0.4032
		Other	National	202	5.3	67.3	1.0	2.9	0.0844	1.41	143.5	0.3210
			State	417	6.1	83.4	1.0	1.0	0.0965	1.37	303.5	0.6790
	Yes	black or Hispanic	National	161	5.2	-	-	-	0.1764	1.35	119.0	0.4378
			State	198	5.0	-	-	-	0.1292	1.30	152.9	0.5622
		Other	National	121	4.8	-	-	-	0.1243	1.29	94.1	0.2502
			State	378	7.0	-	-	-	0.0882	1.34	282.1	0.7498
Indiana	No	black or Hispanic	National	107	5.4	35.7	1.0	3.5	0.3943	1.67	64.1	0.3834
			State	126	3.1	11.5	0.3	1.0	0.1324	1.22	103.1	0.6166
		Other	National	215	5.8	71.7	1.0	3.5	0.0628	1.45	148.1	0.1746
			State	947	5.9	78.9	0.7	1.0	0.1072	1.35	700.1	0.8254
Iowa	No	black or Hispanic	National	2	1.0	2.00	1.0	5.0	0.1837	1.19	1.7	0.0441
			State	45	1.7	5.63	0.8	1.0	0.2007	1.23	36.5	0.9559
		Other	National	146	7.3	73.0	1.0	5.0	0.0997	1.63	89.4	0.1073
			State	1027	6.2	85.6	0.8	1.0	0.1083	1.38	743.7	0.8927
Louisiana	No	black or Hispanic	National	80	4.7	80.0	1.0	5.2	0.2808	1.74	45.9	0.1559
			State	315	3.4	35.0	0.5	1.0	0.1562	1.27	248.4	0.8441
		Other	National	55	4.2	55.0	1.0	5.2	0.1222	1.47	37.5	0.0649
			State	718	5.5	79.8	0.6	1.0	0.1043	1.33	539.9	0.9351
N. Jersey	No	black or Hispanic	National	132	4.3	33.0	1.0	1.0	0.4060	1.57	84.2	0.3293
			State	209	3.4	26.1	0.0	1.0	0.1182	1.22	171.6	0.6708
		Other	National	163	3.5	40.8	1.0	1.0	0.0917	1.23	132.7	0.2375
			State	535	4.6	53.5	0.0	1.0	0.1057	1.26	426.0	0.7625
	Yes	black or Hispanic	National	15	3.0	-	-	-	0.2381	1.32	11.3	0.3438
			State	28	4.7	-	-	-	0.1391	1.29	21.7	0.6562
		Other	National	38	5.4	-	-	-	0.1346	1.32	28.8	0.2554
			State	103	4.9	-	-	-	0.0636	1.23	83.9	0.7446

Table 3-2. Derivation of factors used to composite the National Adult Literacy Survey national and state data — Continued

State	National Certainty PSU	Race/ethnicity	Data source	Sample Size	Persons/ segment	Persons PSU	P_{hk}*	F_h*	Relvariance of weights	Effective design effect	Sample size	Compositing factor
New York	No	black or Hispanic	National	69	5.3	34.5	1.0	0.8	0.2721	1.47	46.9	0.7075
			State	24	1.6	6.00	1.0	1.0	0.2096	1.24	19.4	0.2925
		Other	National	154	5.9	77.0	1.0	0.8	0.1035	1.35	113.8	0.2994
			State	370	6.1	92.5	1.0	1.0	0.1083	1.39	266.3	0.7006
	Yes	black or Hispanic	National	275	7.6	-	-	-	0.3344	1.61	170.5	0.5812
			State	170	7.1	-	-	-	0.1283	1.38	122.9	0.4188
		Other	National	186	5.0	-	-	-	0.2343	1.40	132.5	0.3766
			State	317	9.1	-	-	-	0.1063	1.44	219.4	0.6235
Ohio	No	black or Hispanic	National	158	4.8	31.6	1.0	2.0	0.3722	1.58	100.0	0.4724
			State	138	2.8	11.5	0.2	1.0	0.1579	1.23	111.9	0.5277
		Other	National	309	4.8	61.8	1.0	2.0	0.0962	1.35	229.0	0.2583
			State	871	5.7	72.6	0.4	1.0	0.1040	1.32	657.4	0.7417
Pennsylvania	No	black or Hispanic	National	25	2.3	6.25	1.0	1.4	0.6318	1.69	14.8	0.2555
			State	52	2.5	7.43	0.5	1.0	0.1427	1.21	43.1	0.7445
		Other	National	309	5.9	77.3	1.0	1.4	0.0818	1.37	225.2	0.3048
			State	704	6.2	88.0	0.7	1.0	0.1055	1.37	513.6	0.6952
	Yes	black or Hispanic	National	60	3.5	-	-	-	0.1565	1.26	47.5	0.4881
			State	64	3.4	-	-	-	0.1848	1.28	49.8	0.5119
		Other	National	79	4.2	-	-	-	0.1581	1.29	61.2	0.2693
			State	210	5.3	-	-	-	0.8570	1.26	166.1	0.7308
Texas	No	black or Hispanic	National	235	3.5	58.8	1.0	2.4	0.3547	1.56	150.4	0.4069
			State	272	3.9	34.0	0.9	1.0	0.0942	1.24	219.3	0.5932
		Other	National	250	3.6	62.5	1.0	2.4	0.1670	1.39	180.0	0.3210
			State	497	5.1	62.1	0.8	1.0	0.0971	1.30	380.9	0.6790
	Yes	black or Hispanic	National	194	6.3	-	-	-	0.3709	1.59	121.9	0.5185
			State	145	5.0	-	-	-	0.1132	1.28	113.2	0.4815
		Other	National	155	5.7	-	-	-	0.1429	1.34	115.5	0.3532
			State	320	10.7	-	-	-	0.1068	1.51	211.5	0.6468
Washington	No	black or Hispanic	National	13	1.6	13.0	1.0	1.7	0.4044	1.45	9.0	0.1578
			State	55	1.3	6.88	0.3	1.0	0.1305	1.15	48.8	0.8422
		Other	National	99	6.2	99.0	1.0	1.7	0.0945	1.44	68.8	0.0821
			State	1064	6.5	133.0	0.4	1.0	0.1096	1.38	769.7	0.9179

* As defined in Section 3.2.4.3.

3.2.5 Computing Final Weight—Poststratification Through Raking Ratio Adjustments

Poststratification is commonly used in sample surveys to accomplish three purposes: (1) It generally reduces the sampling errors; (2) it is frequently an effective way of making nonresponse adjustments; and (3) it creates consistency with statistics from other studies. The National Adult Literacy Survey used a particular form of poststratification referred to as raking ratio adjustments. The final sampling weights were computed by raking the composited weights to known population totals. In poststratification, classes are formed from cross-tabulations of certain variables. In some instances, such cross-tabulations may lead to sparse cells, or population distributions may be known for the marginal but not the joint distributions for variables used to define the weighting classes. Weighting class adjustments based on small cell sizes can result in a large amount of variation in the adjusted weights. Raking ratio adjustments are useful for maintaining the weighted marginal distributions of variables used to define weighting classes. For this type of adjustment, population distributions are required for the marginal distributions of the weighting class variables and not for their joint distribution.

An objective of raking ratio adjustments is to adjust the weights of cells in such a way that the marginal distributions for the weighted sample correspond to known population distributions. To illustrate the algorithm, consider a simple case of two variables that are cross-tabulated. Using an example from Kalton (1981), the marginal and joint distributions for the population and sample are as follows.

	Population						Sample				
	1	2	…	K	Total		1	2	…	K	Total
1	W_{11}	W_{12}	…	W_{1K}	$W_{1.}$	1	q_{11}	q_{12}	…	q_{1K}	$q_{1.}$
2	W_{21}	W_{22}	…	W_{2K}	$W_{2.}$	2	q_{21}	q_{22}	…	q_{2K}	$q_{2.}$
⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮		⋮	⋮
H	W_{H1}	W_{H2}	…	W_{HK}	$W_{H.}$	H	q_{H1}	q_{H2}	…	q_{HK}	$q_{H.}$
Total	$W_{.1}$	$W_{.2}$	…	$W_{.K}$	$W_{..}$	Total	$q_{.1}$	$q_{.2}$	…	$q_{.K}$	$q_{..}$

The iterative procedure makes successive modifications to the weights until the process stabilizes. The algorithm used for raking in the National Adult Literacy Survey, and described by Kalton, first weights each cell in row h (h=1, …, H) by the factor $W_{h.}/q_{h.}$. The result is that the sum of the weighted cells for a given row h, $\frac{W_{h}}{q_{h}} \sum_{h} q_{hi}$, will be equal to $W_{h.}$. Because of the adjustments to the weights, the column totals for the sample now become $\sum_{h} q_{hk} \frac{W_{h.}}{q_{h.}} = \sum_{h} q'_{hk} = q'_{.k}$. At the second step in the iterative procedure, the sampled units in each cell in column k (k=1, …, K) are weighted by the factor $W_{.k} \_ q'_{.k}$. Then, the sum of the weights in a given column k is equal to $W_{.k}$. At this point, the $q'_{.k}$ values have been changed to $\sum_{h} q'_{hk} \frac{W_{h.}}{q'_{h.}} = \sum_{h} q''_{hk} = q''_{.k}$. The process now repeats with step one. The procedure is completed when the process converges or, alternatively, is terminated after a pre-specified number of iterations. The result is a set of adjusted weights that are then used for estimation. It has been shown that the raking ratio estimation procedure produces best asymptotically normal estimates under simple random sampling. At the same time, the procedure minimizes the adjustments to the sample weights based on one measure of closeness (Ireland and Kullback, 1968).

Construction of weighting classes is an important consideration in poststratification, particularly when it is used as an adjustment for unit nonresponse. A purpose of using weighting classes is to bring together respondents and nonrespondents with similar characteristics not only for the variables defining the classes but also for variables that are unknown for nonrespondents only. The variables used to construct raking classes for the National Adult Literacy Survey were age, race/ethnicity, sex, education, and geographic indicators, i.e., metropolitan statistical area (MSA) vs. non-MSA for the 11 states and census region for the remainder of the United States.

The 1990 census totals used for raking were adjusted separately by age, race/ethnicity, sex, and region of the country to account for undercoverage. The undercoverage rates used in this process were supplied by the U.S. Bureau of the Census.

3.3 REPLICATED WEIGHTS FOR VARIANCE ESTIMATION IN THE HOUSEHOLD POPULATION

Variance estimation must take into account the sample design. In particular, the estimate of sampling variance for any statistic should account for the effects of clustering, the use of nonresponse and poststratification adjustments, and the component of sampling variability arising from the variation in the weights used to compute the statistic. Treating the data as a simple random sample will produce underestimates of the true sampling variability.

The jackknife method can be used to estimate the variance for most statistics. Jackknifing estimates the sampling variability of any statistic Y, as the sum of components of variability that may be attributed to individual pairs of first-stage sampling units. The variance attributed to a particular pair is measured by estimating how much the value of the statistic would change if only one unit in the pair had been sampled. When using replication techniques such as jackknifing to calculate standard errors, it is necessary to establish a number of subsamples (or replicates) from the full sample, calculate the estimate from each subsample, and sum the squared difference of each replicated estimate from the full-sample estimate. The 60 replicates formed for the National Adult Literacy Survey provided the degrees of freedom necessary for the production of stable estimates of variance.

Variance estimation requires three steps: (1) forming the replicates, (2) constructing the replicate weights, and (3) computing estimates of variance for survey statistics. The formation of replicates is discussed in detail in sections 3.3.1 through 3.3.3. After the replicates had been formed, a replicate factor was constructed for each variance stratum (usually PSUs however, when they are certainty PSUs they are paired). Let $f_{ijk}(r)$ denote the rth replicate factor for the kth respondent in the jth variance unit (member within variance stratum) in the ith variance stratum. Then, in general:

$$f_{ijk}(r) = \begin{cases} 2 \text{ if } i = r \text{ and } j = 1 \\ 0 \text{ if } i = r \text{ and } j = 2 \\ 1 \text{ if } i \neq r \end{cases} \qquad (11)$$

and the replicated base weight, $Wb_{ijk}(r)$, was obtained as $Wb_{ijk}(r) = Wb_{ijk} \, f_{ijk}(r)$ for r = 1, 2, …, 60. (A variation on this scheme, used for only non-certainty PSUs in the state component, is described in section 3.3.2.)

After obtaining a person base weight for each replicate, all remaining full-sample weighting steps leading to the final person weight were performed on each replicate. By repeating the various weight adjustment procedures on each set of replicate base weights, the impact of these procedures on the sampling variance of the estimator Y is appropriately reflected in the variance estimator, v(Y).

After the replicate weights had been constructed, the estimate of variance could easily be computed for any statistic. The statistic was computed 61 times, once using the full-sample weight and an additional 60 times using each of the 60 replicate weights. The variance estimate is the sum of the 60 squared differences between the estimate derived using the full-sample weight and the estimate derived using each of the 60 replicate weights. That is, the estimate of the variance of a statistic Y is,

$$v(Y) = \sum_{r=1}^{60} (Y_r - Y)^2 \qquad (12)$$

where Y_r = the weighted estimate obtained using the rth replicate weight; and Y = the weighted estimate obtained using the full-sample weight.

The National Adult Literacy Survey pooled data from a nationally representative sample of 101 PSUs and from 11 independently selected state PSU samples. The threefold objective of the replication scheme was (1) to reflect the actual sample design of each sample; (2) to ensure the production of stable estimates of standard errors by having sufficient degrees of freedom for national estimates, individual state estimates, and regional estimates; and (3) to limit the total number of replicates so that variance estimation would not be prohibitively expensive. The general approach in setting up the replication was to devise an appropriate scheme for each component of the sample, the national sample and the 11 states, and then to collapse replicates to a reasonable number.

3.3.1 Household Sample Replication for the National Component

The national sample contained 101 PSUs, 25 of which were selected with certainty. The remaining 76 PSUs were selected 2 per stratum using the Durbin method (1967), with probabilities proportional to size and with known joint probabilities. Ordinarily, replicates are formed by pairing first-stage sampled units, that is, segments are paired in PSUs selected with certainty and whole PSUs are paired in non-certainty strata. However, under the Durbin scheme, an unbiased estimate of variance can be obtained by treating PSUs in some non-certainty strata as if they had been chosen with certainty, that is, by pairing segments instead of whole PSUs. For the 101-PSU sample, the natural pairing led to 74 replicates. These replicates were examined carefully to see which contained data from any of the 11 participating states. In certainty PSUs where segments from a participating state had been paired to form a replicate, the segments were grouped into subsets and were paired within each subset to increase the number of replicates and hence the degrees of freedom of the state variance estimator. This procedure expanded the number of national sample replicates to 111.

3.3.2 Household Sample Replication for the State Component

An independent sample of 8 to 12 PSUs was selected in each of the 11 participating states. The largest PSUs were taken with certainty. Within each state, the remaining PSUs were grouped into strata, and from each stratum a single PSU was sampled with probability proportional to size. In PSUs selected with certainty, segments were paired to form replicates. However, the segments were grouped into subsets and paired within each subset to increase the degrees of freedom. This procedure created from 2 to 8 replicates for each PSU chosen with certainty, with a total of 113 replicates across the 11 states.

Ordinarily, non-certainty PSUs would be paired to form replicates so that, for instance, a state with *n* such PSUs would yield *n*/2 replicate pairs. With the goal of increasing the degrees of freedom, an alternative procedure was adopted. The same *n* PSUs were used to create *n*-1 replicates, as follows: The active part of each replicate contained data from exactly *n*-1 of the *n* PSUs, and the base weight was multiplied by $n/(n-1)$ rather than the usual factor of 2. One randomly selected PSU was active[3] in all n-1 replicates, and a successively different one of the remaining n-1 PSUs was inactive[4] in each of the n-1 replicates. The applied method kept estimation consistent with the rest of the sample and created 54 replicates across the 11 states.

3.3.3 Final Household Sample Replication for the National and State Components

A total of 278 replicates had been formed at this point: 111 from the national sample, 113 from PSUs chosen with certainty for the state samples, and 54 from non-certainty PSUs chosen for the state samples. These replicates reflected the actual design of each sample and provided sufficient degrees of freedom to produce stable estimates of variance for the nation, each state, and the four census regions. However, using 278 replicates to estimate variances would be computer intensive and expensive, while providing only a slight gain in the precision of the overall estimates. Therefore, the replicates were collapsed to 60, a much more realistic number. To preserve the total number of replicates for each state, replicates from the same state were never collapsed. As often as possible, the same constraint was used by region as well.

Table 3-3 presents the results of the replication scheme, showing which replicates are active for the major sub-domains of analysis.

3.4 CALCULATING SAMPLE WEIGHTS FOR THE PRISON POPULATION

The final inmate weight was constructed in four major steps. The first step was to construct the inmate base weight, which was the reciprocal of the overall probability of selection for each inmate. The second step was to adjust the inmate base weight for the one facility that did not cooperate, so that weighted estimates for inmates from cooperating facilities would also represent inmates from the non-cooperating facility. The third step was to adjust the inmate weight to compensate for not obtaining a completed background questionnaire for every inmate in the sample. The fourth step was to post-stratify the weight so that the weighted counts from the sample agreed with independent estimates for certain subgroups of the population.

3.4.1 Computing Inmate Base Weights

The initial correctional facility sample consisted of 96 facilities, of which eight facilities were randomly selected and set aside as the reserve sample. The reserve sample was never used because the actual response rates were higher than those originally estimated for the sample of 96 facilities. The reduced sample of facilities was drawn by taking a systematic sample, with equal probabilities of selection, from a listing of all sample facilities in their initial selection order.

---

[3] An active PSU is one that contributes to the replicates' estimate.

[4] An inactive PSU does not contribute to the replicates' esitmate.

$$W_{bi} = \frac{1}{P_i} \frac{1}{88 \ / 96} \quad (13)$$

The facility weight for the remaining 88 facilities in the sample was computed as a product of the reciprocal of the probability of the i^{th} facility (PSU) being selected to the initial sample and the reciprocal of the probability of its not being selected to the reserved sample; that is:

where

W_{bi} = the base weight for the i^{th} facility; and

P_i = the probability of selection of the i^{th} facility.

The inmate base weight is the reciprocal of the overall probability of selecting the j^{th} inmate in the i^{th} facility.

$$WI_{bij} = W_{bi} \frac{N_i}{n_i} \quad (14)$$

where

N_i = the inmate population size for the i^{th} facility; and

n_i = the inmate sample size for the i^{th} facility.

WI = weight for the base weight for j^{th} inmate in the i^{th} facility.

Table 3-3. Active replicates for sub-domains of the National Adult Literacy Survey analysis file

Replicate	Household sample																Prison sample
	U.S.	Northeast	Midwest	South	West	California	Illinois	Indiana	Iowa	Louisiana	New Jersey	New York	Ohio	Pennsylvania	Texas	Washington	
1	x	x	x	x	x	x	x		x	x	x						x
2	x	x	x	x	x	x			x	x	x						x
3	x	x	x	x	x	x	x		x	x	x		x				x
4	x	x	x	x	x	x	x	x	x	x	x						x
5	x	x	x	x	x	x	x		x	x	x						x
6	x	x	x	x	x	x	x		x	x	x						x
7	x	x	x	x	x	x	x		x	x	x						x
8	x	x	x	x	x	x	x		x	x	x						x
9	x	x	x	x	x	x	x		x	x	x						x
10	x	x	x	x	x	x	x		x	x	x						x
11	x	x	x	x	x	x	x		x	x	x						x
12	x	x	x	x	x	x	x		x	x	x						x
13	x	x	x	x	x	x	x		x	x	x						x
14	x	x	x	x	x	x	x		x	x	x						x
15	x	x	x	x	x	x	x		x	x	x						x
16	x	x	x	x	x	x	x		x	x	x						x
17	x	x	x	x	x	x	x		x	x	x						x
18	x	x	x	x	x	x	x		x	x	x						x
19	x	x	x	x	x	x	x		x	x	x						x
20	x	x	x	x	x	x	x		x	x	x						x
21	x	x	x	x	x	x	x				x						x
22	x	x	x	x	x	x	x				x						x
23	x	x	x	x	x	x	x				x						x
24	x	x	x	x	x	x	x				x						x
25	x	x	x	x	x	x		x			x						x
26	x	x	x	x	x	x		x			x						x
27	x	x	x	x	x	x		x			x						x
28	x	x	x	x	x	x		x			x						x
29	x	x	x	x	x	x		x			x						x
30	x	x	x	x	x	x		x				x					x
31	x	x	x	x	x	x		x				x					x

Table 3-3. Active replicates for sub-domains of the National Adult Literacy Survey analysis file — Continued

Replicate	Household sample																Prison sample
	U.S.	Northeast	Midwest	South	West	California	Illinois	Indiana	Iowa	Louisiana	New Jersey	New York	Ohio	Pennsylvania	Texas	Washington	
32	x	x	x	x	x	x		x				x	x				x
33	x	x	x	x			x				x	x					x
34	x	x	x	x			x				x	x					x
35	x	x	x	x			x				x	x					x
36	x	x	x	x	x			x				x	x		x		x
37	x	x	x	x	x			x				x	x		x		x
38	x	x	x	x	x			x				x	x	x	x		x
39	x	x	x	x	x			x				x	x	x	x		x
40	x	x	x	x	x			x				x	x	x	x		x
41	x	x	x	x	x			x				x	x	X	x	x	x
42	x	x	x	x	x							x	x	X	x	x	x
43	x	x	x	x	x							x	x	x	x	x	x
44	x	x	x	x	x							x	x	x	x	x	x
45	x	x	x	x	x							x	x	x	x	x	x
46	x	x	x	x	x							x	x	x	x	x	
47	x	x	x	x	x							x	x	x	x	x	
48	x	x	x	x	x							x	x	x	x	x	
49	x	x	x	x	x							x	x	x	x	x	
50	x	x	x	x	x							x	x	x	x	x	
51	x	x	x	x	x							x	x	x	x	x	
52	x	x	x	x	x								x	x	x	x	
53	x	x	x	x	x								x	x	x	x	
54	x	x	x	x	x								x	x	x	x	
55	x	x	x	x	x								x	x	x	x	
56	x	x	x	x	x									x	x	x	
57	x	x	x	x	x									x	x	x	
58	x	x	x		x											x	
59	x	x	x		x											x	
60	x	x	x														
# active	60	60	60	57	59	32	23	18	20	20	29	22	25	20	22	19	45

77

78

3.4.2 Nonresponse Adjustments

3.4.2.1 Facility nonresponse adjustment

Only one correctional facility did not cooperate. As described in section 2.6.1.1, the sample facilities were stratified on the basis of certain characteristics. Using this stratification scheme, the non-cooperating facility was classified as a state maximum security facility, in the southern region of the United States, with a male-only inmate population. To adjust for the non-cooperating facility, two nonresponse adjustment classes were constructed: (1) all facilities in the same sampling stratum (implicit stratum) as the non-cooperating facility and (2) all remaining facilities. The facility nonresponse adjustment factor was computed for each nonresponse class as the ratio of the weighted (facility weight times the facility inmate population size) sum of all eligible sample facilities to the respondent facilities. That is, the nonresponse adjustment factor for the α^{th} class, $A_{F\alpha}$, was computed as

$$A_{F\alpha} = \frac{\sum_{i \in S(\alpha)} W_{b\alpha_i} N_{\alpha_i}}{\sum_{i \in SR(\alpha)} W_{b\alpha_i} N_{\alpha_i}} \tag{15}$$

where

$W_{b\alpha i}$ = the facility weight for the i^{th} facility in the α^{th} facility nonresponse adjustment class;

$N_{\alpha i}$ = the inmate population count for the i^{th} facility in the α^{th} facility nonresponse adjustment class;

$S(\alpha)$ = the collection of all eligible (cooperating and non-cooperating) sample facilities in the α^{th} facility nonresponse adjustment class; and

$SR(\alpha)$ = the collection of all cooperating facilities in the α^{th} facility nonresponse adjustment class.

Table 3-4 presents the facility nonresponse adjustment factors for both nonresponse adjustment classes.

Table 3-4. The National Adult Literacy Survey correctional facility sample counts and facility nonresponse adjustment factor, by facility nonresponse adjustment classes

Nonresponse adjustment class	Sample count		Nonresponse adjustment factor
	Eligible	Respondent	
1	8	7	1.122
2	80	80	1.000

3.4.2.2 Inmate nonresponse adjustment

The inmate sample consisted of 1,340 inmates, of whom 1,147 completed background questionnaires. The main reason for adjusting the sampling weights was to remove potential bias on statistics of interest as a result of the inability to collect completed background questionnaires for all sample inmates. If the probability of nonresponse were independent of the statistics of interest, then no bias would arise. Therefore, the objective was to obtain adjustment classes such that the probability of nonresponse within each class was as independent of statistics of interest as possible. There are several alternative methods of forming the classes to achieve this result. For the prison sample, the classes were formed so that the variation in the response propensity within the classes was minimized.

A set of potential predictive variables was selected for the response propensity. These variables had to be available for respondents and nonrespondents alike. They were

- State vs. Federal facility;
- Region: Northeast, Midwest, South, West;
- Sex of inmates: male only, both sexes, female only; and
- Facility type: maximum security, medium security, minimum security, medical, all other.

To form the nonresponse adjustment classes, a technique similar to the automatic interaction detection type of algorithm was used. Pearson chi-square statistics were computed between the response and each one of the predictive variables. The predictor with the smallest p-value was selected as the "best" predictor. Then, the same process was applied within the subgroups of the population, defined by the levels of the "best" predictor chosen in the preceding step. This process was continued until no significant predictor was found or until a specified minimum class size had been reached. The procedure is stepwise and creates a hierarchical, tree-like structure. The inmate nonresponse classes are shown in Table 3-5.

Table 3-5. The National Adult Literacy Survey inmate sample counts and nonresponse adjustment factors, by inmate nonresponse adjustment classes

Region	Facility type	State/Federal facility	Sample counts		Nonresponse adjustment factor
			All	Respondent	
Northeast and West	Maximum security and medical	All	171	121	1.386
Northeast and West	All other	State	330	275	1.196
Northeast and West	All other	Federal	54	51	1.063
South and Midwest	Maximum security and medical	All	212	174	1.214
South and Midwest	Medium security	All	337*	302*	1.117
South and Midwest	Minimum security and other	All	235	224	1.051

*This class actually contained 338 and 303 responding inmates, with the additional unit representing one inmate who was selected into the sample twice from two different facilities. The number of records is adjusted here to be consistent with the number of records (1,147) receiving weights.

The inmate nonresponse adjustment factor for the hth nonresponse adjustment class, INRAF$_h$, was computed as

$$\mathrm{AIh} = \frac{\sum_{i \in A(h)} WI_{bhi} A_{Fhi}}{\sum_{i \in AR(h)} WI_{bhi} A_{Fhi}} \tag{16}$$

where

WI_{bhi} = the base weight for the ith inmate in the hth inmate nonresponse adjustment class;

A_{Fhi} = the facility nonresponse adjustment factor for the ith inmate in the hth nonresponse adjustment class;

A(h) = the collection of all sample inmates in the hth facility nonresponse adjustment class; and

AR(h) = the collection of all sample inmates with completed background questionnaires in the hth facility nonresponse adjustment class.

3.4.3 Poststratification Procedures

To reduce the mean square error of estimates, the weights were further adjusted so that the weighted totals obtained from the sample as estimates for certain subgroups of the population would be consistent with presumably more precise estimates available from external sources. Control totals were obtained from the U.S. Department of Justice's Bureau of Justice Statistics and were partly based on data from the 1991 Survey of Inmates in State Correctional Facilities. Both sets of estimates were obtained from larger samples than the one utilized in this survey and thus were expected to have greater precision.

Poststratification was intended to reduce nonresponse-related residual bias on the estimates and simultaneously to increase the precision of the post-stratified estimates. This beneficial effect on the variance was not restricted to the post-stratified variables. The precision of any substantive variable correlated with the post-stratified variables was also expected to improve.

For the male inmates, the poststratification estimation utilized raking ratio estimation. The inmate nonresponse adjusted weights were alternately adjusted by an iterative process to provide consistency with the independent estimates of population by age and then by education within each race/ethnicity category. Table 3-6 shows the sample estimates for male inmates (before raking) and the independent control totals by age and by education within race/ethnicity categories.

Table 3-6. Comparison of the National Adult Literacy Survey sample estimates (before raking) and independent control totals, by age and by education within ethnicity, for male inmates

Race/ethnicity	Age or education	Sample		Control total
		Size	Estimate	
White and other	Age			
	less than 30	173	117,604	107,332
	30 or more	255	175,019	167,488
	Education			
	0-8 years	49	34,375	31,496
	9-12 years	272	186,001	189,149
	Some college	107	72,246	54,175
black	Age			
	less than 30	240	165,229	155,912
	30 or more	210	145,130	164,931
	Education			
	0-8 years	40	27,475	35,968
	9-12 years	333	230,590	239,645
	Some college	77	52,118	45,230
Hispanic	Age			
	less than 30	107	76,144	6,400
	30 or more	91	61,543	65,569
	Education			
	0-8 years	59	41,256	34,035
	9-12 years	109	76,144	77,758
	Some college	30	20,289	15,176

Raking ratio estimation was used rather than a straightforward poststratification procedure because the cell sizes were too small to obtain stable estimates when age and education were cross-classified within race/ethnicity. Refer to section 3.2.5 for a detailed description of raking ratio estimation.

Table 3-7 shows the raking ratio estimate and the adjustment factor for each adjustment class for the male inmates. The small adjustment factors for inmates with some college education could be related to the tendency of better educated inmates to be more cooperative. A similar pattern can be observed for inmates who were less than 30 years old.

Table 3-7. Raking ratio estimates and weight adjustment factors for male inmates in the National Adult Literacy Survey sample

Adjustment cell	Race/ethnicity	Education	Age	Adjustment factor
1	White and other	0-8 years	less than 30	0.859
2			30 or more	0.943
3		9-12 years	less than 30	0.067
4			30 or more	1.061
5		Some college	less than 30	0.700
6			30 or more	0.768
7	black	0-8 years	less than 30	1.130
8			30 or more	1.397
9		9-12 years	less than 30	0.952
10			30 or more	1.177
11		Some college	less than 30	0.741
12			30 or more	0.910
13	Hispanic	0-8 years	less than 30	0.684
14			30 or more	0.950
15		9-12 years	less than 30	0.892
16			30 or more	1.239
17		Some college	less than 30	0.614
18			30 or more	0.853

One-dimensional poststratification was used for female inmates mainly because of the small sample size for this group. The poststratification adjustment factor for the g^th^ poststratification adjustment class, PA_{Ig}, was

$$PA_{Ig} = \frac{C_g}{\sum_{i \in E(g)} WI_{bgi} A_{Fgi} A_{Igi}} \qquad (17)$$

where

C_g = the female inmate control total for the g^th^ poststratification class;

$E(g)$ = the collection of female respondent inmates in the g^th^ poststratification class;

WI_{bgi} = the inmate base weight for the i^th^ inmate in the g^th^ poststratification class;

A_{Fgi} = the facility nonresponse adjustment factor for the i^th^ inmate in the g^th^ poststratification class; and

A_{Igi} = the inmate nonresponse adjustment factor for the i^th^ inmate in the g^th^ poststratification class.

The poststratification factors for the female inmates are shown in Table 3-8.

Table 3-8. Control totals and poststratification adjustment factors for female inmates in the National Adult Literacy Survey sample, by poststratification classes

Poststratification adjustment cell	Race/ethnicity	Sample size	Control total	Poststratification factor
19	black	30	19,465	0.906
20	All other	41	23,554	0.875

3.4.4 Final Inmate Weights

Final inmate weights were obtained as a product of the inmate base weight, the facility nonresponse adjustment factor, the inmate nonresponse adjustment factor, and the raking/poststratification adjustment factor:

$$Fw_{gh\alpha i} \quad = \quad WI_{bgh\alpha i}\ A_{F\alpha i}\ A_{Ihi}\ PA_{Igi} \qquad (18)$$

where

$WI_{bgh\alpha i}$ = the base weight for the ith inmate in the α^{th} facility nonresponse adjustment class, the hth inmate nonresponse adjustment class, and the gth poststratification class;

$A_{F\alpha i}$ = the facility nonresponse adjustment factor for the ith inmate in the α^{th} facility nonresponse adjustment class;

A_{Ihi} = the inmate nonresponse adjustment factor for the ith inmate in the hth inmate nonresponse adjustment cell; and

PA_{Igi} = the poststratification/raking adjustment factor for the ith inmate in the gth poststratification class.

Table 3-9 presents statistics for the sampling weights at each stage of weight adjustment. The table shows that the variation in the base weight was rather small and that nonresponse adjustments had only a trivial effect on the weight variation. The poststratification/raking increased the weight variation moderately. Despite the increase in weight variation, poststratification/raking usually decreases the variance of estimates for any characteristics that are correlated with the raked variables (Brackstone and Rao, 1979; Oh and Scheuren, 1978). The post-stratified/raked variables in this survey are known to be strongly correlated with many substantive characteristics. The poststratification procedure was effective in simultaneously reducing the residual nonresponse bias and the sampling variance.

Table 3-9. Statistics for the distribution of the weight-by-weight adjustment stage for the National Adult Literacy Survey incarcerated sample

Statistic	Base weight	Facility nonresponse adjusted weight	Inmate nonresponse adjusted weight	Post-stratified raked weight
Sample size	1,340	1,340	1,147	1,147
Mean	582.52	588.16	687.13	667.52
cv (%)	16.51	16.56	18.43	24.94
Minimum	110.22	110.22	115.89	110.29
5th Percentile	491.47	491.47	530.58	458.49
Median	593.20	596.27	684.30	644.83
95th Percentile	680.51	700.67	877.48	937.89
Maximum	1,012.37	1,012.37	1,682.92	1,785.87

3.5 REPLICATED WEIGHTS FOR VARIANCE ESTIMATION IN THE PRISON POPULATION

The use of a complex sample design, adjustments for nonresponse, and poststratification procedures resulted in dependence among the observations. The application of the usual formulae of variance estimation, which were based on simple random sampling assumptions, would result in the underestimation of sampling variance in this survey. To estimate sampling variability, therefore, 45 jackknife replicates were formed to provide adequate degrees of freedom for the production of reliable estimates. The variance estimation was carried out in three steps: (1) the replicates were formed, (2) the replicate weights were computed, and (3) the estimates of the variances of the survey statistics were computed.

The replicates were designed in accordance with the sample design. The 86 non-certainty facilities were placed in their sample selection order. Then, the facilities were paired consecutively, and each pair was assigned to a variance stratum. This process resulted in 43 variance strata. Within each variance stratum, one facility was assigned randomly to variance unit 1 and the other to variance unit 2. The two largest facilities in the sample were assigned to separate variance strata. These facilities were certainty selections and therefore their only contribution to the total variance was from within-facility sampling. Therefore, the inmate records within each facility were placed in their sample selection order and numbered sequentially. The odd-numbered inmates were assigned to one variance unit and the even-numbered inmates to the other. Thus, a total of 45 variance strata and 90 variance units were obtained. After the replicates had been formed, the replicate weights were constructed. A replicate factor was constructed for each variance stratum. If $f_{ijk}(r)$ denotes the r^{th} replicate factor for the k^{th} inmate in the j^{th} variance unit and the i^{th} variance stratum, then

$$f_{ijk}(r) = \begin{cases} 2 \text{ if } i = r \text{ and } j = 1 \\ 0 \text{ if } i = r \text{ and } j = 2 \\ 1 \text{ if } i \neq r \end{cases} \qquad (19)$$

The rth replicate inmate base weight for the kth inmate in the ith variance stratum and the jth variance unit, $WI_{hijk}(r)$, was then obtained as

$$WI_{hijk}(r) \qquad = \qquad WI_{hijk}\ f_{ijk}(r) \tag{20}$$

for r = 1, 2,........45.

After obtaining an inmate base weight for each replicate, all remaining full-sample weighting steps leading to the final inmate weight were performed on each replicate. For each replicate, a facility nonresponse adjustment factor, an inmate nonresponse adjustment factor, and a poststratification adjustment factor were computed, and these factors were then applied to the replicate inmate base weight to obtain 45 replicate final inmate weights. Replicate weights 46 through 60 were "inactive" for the prison sample and were set equal to the full-sample weight in the data file. The variance estimation procedures were similar to those used for the household sample, as described in section 3.3.

Chapter 4

DEVELOPMENT OF THE SURVEY INSTRUMENTS

Anne Campbell, Diné College (formerly of Educational Testing Service)

One of the goals of the 1992 National Adult Literacy Survey was to relate the literacy skills of the nation's adults to a variety of demographic characteristics and explanatory variables. To accomplish this goal, the survey included the administration of a background questionnaire as well as literacy simulation tasks. The next three sections describe the conceptual framework for the survey and the development of the background questionnaire and the literacy tasks.

4.1 CONCEPTUAL FRAMEWORK

One of the major goals of the National Adult Literacy Survey (NALS) was to compare its results with those from other large-scale assessments of literacy that have been conducted during the past few years. These include two major surveys: 1) the 1985 Young Adult Literacy Assessment, conducted as a part of the National Assessment of Educational Progress (NAEP) and carried out by Educational Testing Service (ETS) and the Response Analysis Corporation under a grant from the National Center for Education Statistics (NCES; Kirsch and Jungeblut, 1986), and 2) the 1990 Workplace Literacy Survey, conducted by ETS under a contract from the Employment and Training Administration (Kirsch, Jungeblut, and Campbell, 1992). Thus, the conceptual framework for the National Adult Literacy Survey is based on the framework developed for the Young Adult Literacy Assessment and used again in the Workplace Literacy Survey.

The foundation for the 1985 Young Adult Literacy Assessment, the 1990 Workplace Literacy Survey, and the 1992 National Adult Literacy Survey was the following definition of literacy:

> *Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential.*

This definition characterizes literacy by focusing on what adults do with printed and written information. It rejects an arbitrary standard, such as signing one's name, completing five years of schooling, or scoring at the eighth grade level on a test of reading achievement. In addition, this definition goes beyond simply decoding and comprehending text and implies that the information-processing skills that adults use to think about content are part of the concept of literacy.

The National Center for Education Statistics specified in its contract requirements for conducting the National Adult Literacy Survey that ETS appoint a Literacy Definition Committee to provide substantive expertise to guide the development and conduct of the survey. The Literacy Definition Committee recommended adopting the above definition of literacy, along with the three literacy scales developed to report the results of the Young Adult Literacy Assessment as the framework for the National Adult Literacy Survey.

Three literacy scales—prose literacy, document literacy, and quantitative literacy—were also used in the two preceding national surveys of literacy and represent distinct and important aspects of the ability to use printed and written information.

Prose literacy consists of the knowledge and skills needed to understand and use information contained in prose texts, both expository and narrative. Expository prose consists of printed information in the form of connected sentences and longer passages that define, describe, or inform, such as newspaper stories or written instructions. Narrative prose tells a story, but is less frequently used by adults in everyday life than by school children, and did not occur as often in the texts presented in the prose literacy tasks. Prose varies in its length, density, and structure (e.g., use of section headings or topic sentences for paragraphs). Using information contained in prose texts, or prose literacy, means that people can locate information contained in prose in the presence of related, but unnecessary information, find all the information, integrate information from various parts of a passage of text, and write new information related to the text.

Document literacy consists of the knowledge and skills required to locate and use information found in documents. Documents differ from prose text in that they are more highly structured. Documents consist of structured prose and quantitative information, in complex arrays arranged in rows and columns, such as tables, data forms, and lists (simple, nested, intersected, or combined), in hierarchical structures such as tables of contents or indexes, or in two-dimensional visual displays of quantitative information, such as graphs, charts, and maps. Using information contained in documents, or document literacy, means that people can locate information in documents, repeat the search as many times as needed to find all the information, integrate information from various parts of a document, and write new information as requested in appropriate places in a document, while screening out related, but inappropriate information.

Quantitative literacy consists of the knowledge and skills needed to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials. Quantities can be located in either prose texts or in documents. Quantitative information may be displayed in analog form in graphs, maps, or charts, or it may be displayed in digital form using whole numbers, fractions, decimals, percentages, or time units (hours and minutes). Using quantitative information contained in prose or documents, or quantitative literacy, means that people can locate quantities while screening out related, but unneeded information, repeat the search as many times as needed to find all the numbers, integrate information from various parts of a text or document, infer the necessary arithmetic operation(s), and perform the arithmetic operation(s) correctly.

The three literacy scales were measured with literacy tasks that simulate the demands that adults encounter when they interact with printed materials on a daily basis (simulation tasks). The tasks used to measure literacy along the three scales incorporate many features designed to demonstrate that adults can use information, including quantitative information, contained in texts and documents.

The adoption of the definition of literacy and the three scales from the Young Adult Literacy Assessment facilitated implementing the goal of comparing the demonstrated literacy proficiencies of the national survey population with those of the populations from the two prior surveys. To ensure that valid comparisons could be made by linking the scales, a set of 85 tasks that were administered in the Young Adult Literacy Assessment and in the Workplace Literacy Survey were also planned to be included in the 1992 National Adult Literacy Survey. Still, new tasks needed to be developed because some of the old tasks had become dated and because a better balance of tasks among the three scales was needed (about two-thirds of the original tasks contributed to the document scale, leaving one-sixth of the tasks for the prose scale and one-sixth for the quantitative scale).

Taking into consideration the definition of literacy and the three literacy scales, the Literacy Definition Committee established the following guidelines for developing new literacy tasks:

- Continued use of open-ended simulation tasks rather than multiple-choice questions;
- Continued emphasis on measuring a broad range of information-processing skills covering a variety of contexts;
- Increased emphasis on simulation tasks that require brief written and/or oral responses;
- Increased emphasis on tasks that focus on asking the respondent to describe how he or she would set up and solve the problem; and
- The use of a simple, four-function calculator to solve quantitative problems.

Using these guidelines, an additional 81 tasks were developed specifically for the 1992 National Adult Literacy Survey in order to complement and enhance the original set of 85 literacy tasks.

In addition to the definition of literacy and the three literacy scales, the administration of a background questionnaire to collect demographic and background information was also carried over from the 1985 and 1990 assessments. This information, along with the information gathered from the simulation tasks, is important for interpreting and reporting the literacy results.

4.2 THE SCOPE OF THE BACKGROUND QUESTIONNAIRE

The questionnaire was intended to provide data about the U.S. adult population, enhance understanding of the factors related to the observed distribution of literacy skills, and facilitate comparisons with previous studies. A modified version of the questionnaire was developed for the prison population, as some of the questions for the population at large were not relevant for this subgroup (see Appendix H). Both background questionnaires, but not the literacy tasks, were also translated into Spanish.

Two goals guided the development of the questionnaire:

- To ensure the usefulness of the data by addressing issues of concern throughout the nation; and
- To ensure comparability with the Young Adult Literacy Assessment and the Department of Labor Workplace Literacy Survey by including some identical questions.

In keeping with these goals, the background questionnaire addressed the following broad issues:

- General and language background;
- Educational background and experiences;
- Political and social participation;
- Labor force participation;
- Literacy activities and collaboration; and
- Demographic information.

4.2.1 General and Language Background

By design, the survey is a study of English literacy proficiency. Projected demographic changes, however, point to a large and growing population of adults with limited English proficiency. It was likely, therefore, that little or no information from the simulation tasks in English would be available for these individuals and, thus, they could be characterized only from the information collected in the background questionnaire. In addition, many of the questions included in the category of general and language background were important in characterizing the sample of young adults in the 1985 Young Adult Literacy Assessment; and, in fact, the age at which English was learned was found to be a powerful variable in previous analyses of the data on young adults. In order to gather as much pertinent information as possible, the questions relating to respondents' general and language background addressed the following:

- Country of birth;
- Education before coming to the United States;
- Language(s) spoken by others in the home;
- Language(s) spoken while growing up;
- Language(s) spoken now;
- Participation in courses for English as a second language; and
- Self-evaluation of proficiency in English and other languages.

4.2.2 Educational Background and Experiences

Although "self-educated" individuals can still be found, formal education remains among the most important factors in the acquisition of literacy skills. Level of education is known to be an important predictor of demonstrated performance on the prose, document, and quantitative literacy scales across racial/ethnic groups. The questions addressing educational background and experiences were designed to provide data for descriptive and relational analyses as well as to address some specific issues. The questions collected information on the following:

- Highest grade or level of education completed;
- Reasons for not completing high school;
- High school equivalency;
- Current educational aspirations;
- Types and duration of training received in addition to traditional school;
- Context, that is, school, home, or work, in which literacy activities were learned; and
- Physical, mental, or health conditions that may affect literacy skills.

4.2.3 Political and Social Participation

People need to read, write, and calculate in order to accomplish important tasks not only at work and in school, but also at home and in their communities. The questions included under political and social participation make it possible to explore the kinds of free-time activities that adults engage in relative to demonstrated proficiencies. Information on the use of library services is important because libraries promote reading and often provide literacy programs. In addition, because an informed citizenry is essential to political participation, and because printed material is an important medium for conveying information on public issues, information was collected on how adults keep abreast of current events and public affairs. The questions in this section addressed the following:

- Sources for obtaining information about current affairs;
- Television viewing;
- Use of library services; and
- Voting behavior.

4.2.4 Labor Force Participation

There is widespread concern that the literacy skills of both our present and future work forces are not adequate for competing in the current global economy or for coping with our rapidly evolving technological society. The questions relating to labor force participation are based on standard labor force concepts widely used in economic surveys; they allow a variety of labor market activity and experience variables to be constructed. Combined with the data on the demonstrated literacy proficiencies of adults, the labor market variables make it possible to examine associations between literacy proficiencies and the labor market experiences of key subgroups. In addition, the questions included make it possible to link results to the Department of Labor literacy survey. The questions in this section addressed the following:

- Employment status;
- Weekly wages or salary;
- Weeks of employment for the last year;
- Annual wages or salary; and
- Industry and occupation.

4.2.5 Literacy Activities and Collaboration

Questions relating to literacy activities and collaboration addressed several important issues. Some of the questions provided information about the types of materials—newspapers, magazines, books, and brief documents—that adults read, making it possible to investigate the relationship between the types of materials read and demonstrated literacy proficiencies. Another subset of questions asked about the frequency of particular reading, writing, and mathematics activities engaged in for personal use as well as for use on the job. By asking adults about the types of literacy practices they engage in specifically for work, analyses can relate on-the-job literacy practices to various occupational categories, education levels, and income levels. The issue of collaboration was addressed by questions that asked if a person received assistance when engaging in particular literacy activities. The questions in this section collected information on the following:

- Newspaper, magazine, and book reading practices;
- Reading, writing, and mathematics activities engaged in for personal use;
- Reading, writing, and mathematics activities engaged in for work; and
- Assistance received from others with particular literacy activities.

4.2.6 Demographic Information

The inclusion of demographic variables makes it possible to describe the adult population as well as to investigate the demonstrated literacy proficiencies of major subgroups of interest, such as racial/ethnic groups, males and females, and age groups, including those over the age of 64. In addition, the data allow for the investigation of such issues as the educational experiences of White, black, and Hispanic populations as well as their access to literacy related services; the educational experiences of different generations of adults; and the relationships of socioeconomic status and family background to literacy.

The demographic information collected included the following:

- Educational attainment of parents;
- Marital status;
- Number of people in family employed full time and part time;
- Sources of income other than employment;
- Family and personal income from all sources;
- Race/ethnicity;
- Age; and
- Sex.

4.2.7 Prison Survey Background Questionnaire

Because many of the questions for the household population were not appropriate for a prison population, a more relevant version of the background questionnaire was developed incorporating questions from the 1991 Survey of Inmates of State Correctional Facilities, sponsored by the Bureau of Justice Statistics of the U.S. Department of Justice (see Appendix H).

Most of the questions in the household survey questionnaire that dealt with general and language background and with literacy activities and collaboration remained in the incarcerated questionnaire. Many of the questions dealing with education, however, were either revised or replaced with questions from the 1991 inmate survey. These questions better reflected the educational experiences of inmates both prior to their incarceration and while in prison. The questions pertaining to political and social participation in the household questionnaire were replaced with questions from the 1991 inmate survey dealing with current offenses and criminal history. Some of the questions in the household questionnaire dealing with labor force participation were replaced with questions about inmates' prison work assignments. Several questions dealing with family income and employment status of family members were dropped from the demographic section of the questionnaire. As a result of these changes, the questionnaire for the prison population addressed the following major topics:

- General and language background;
- Educational background and experiences;
- Current offenses and criminal history;
- Prison work assignments and labor force participation prior to incarceration;
- Literacy activities and collaboration; and
- Demographic information.

4.2.8 Spanish Versions of the Questionnaires

Because Spanish is the second most prevalent language in this country, both the household and prison background questionnaires were translated into Spanish and administered by bilingual interviewers. The non-English, non-Spanish language groups are not prevalent enough across the country as a whole to make other translations practical for conducting the survey. Because native Spanish speakers may not be able to complete the assessment's simulation tasks in English, it was considered important to collect background information in order to understand the language background and literacy experiences of that group. Since the survey was intended to assess only the English literacy skills of the population, the simulation tasks were not offered in Spanish.

4.3 DEVELOPMENT OF THE SIMULATION TASKS

This section describes the development of the new National Adult Literacy Survey tasks as well as the scope of the combined pool of existing tasks—that is, the original tasks plus the tasks newly developed for the National Adult Literacy Survey. It also describes the process of grouping the tasks into blocks or sections and then assembling these blocks into booklets for administration.

4.3.1 Organizing Framework for Task Development

The framework used to develop the National Adult Literacy Survey tasks reflects research conducted on the tasks from the 1985 Young Adult Literacy Assessment, particularly with respect to the processes and strategies involved in completing the tasks. Thus, the National Adult Literacy Survey tasks served to refine and extend the three existing literacy scales—prose, document, and quantitative literacy.

In developing the tasks for the National Adult Literacy Survey, one goal was to complement the tasks that had been developed for the Young Adult Assessment. This meant including a diversity of stimulus materials and designing tasks that represented the broad range of skills and processes inherent in the three domains of literacy. Furthermore, the tasks were designed to assess a wide variety of skills reflecting the demands adults encounter in occupational, community, and home settings—skills that involve reading, writing, and computing. Because the tasks were meant to simulate the kinds of activities that people engage in when they use printed materials, they were open-ended. The underlying principle for the development of the National Adult Literacy Survey tasks was that demonstrated performance on any given task reflects interactions among the following:

- The structure of the stimulus material, e.g., exposition, narrative, table, graph, map, or advertisement;
- The content represented and/or the context from which the stimulus is drawn, e.g., work, home, community; and
- The nature of what the individual is asked to do with the material, i.e., the purpose for using the material, which guides the strategies needed to complete the task successfully.

4.3.2 Materials/Structures

The stimulus materials selected for the tasks included a variety of structures or linguistic formats that adults encounter in their daily activities. The materials were reproduced in their original format. Most of the prose materials used in the survey were expository—that is, they describe, define, or inform—since much of the prose that people read is expository in nature; however, narratives and poetry were included as well. The expository materials included a diversity of linguistic structures, from texts that were highly organized both topically and visually to those that were loosely organized. They also included texts of varying lengths, from full-page magazine articles to short newspaper articles of several paragraphs.

The document tasks were based on a wide variety of document structures, which were categorized as tables, charts and graphs, forms, maps, and miscellaneous documents. Tables included matrix documents in which information is arrayed in rows and/or columns, such as transportation schedules and lists or tables of information. Documents categorized as charts and graphs included pie charts, bar graphs, and line graphs. Forms included any documents that required information to be filled in, and miscellaneous structures included such materials as advertisements and coupons.

Because quantitative tasks involve performing arithmetic operations on numbers embedded in print, they were based on some kind of stimulus material. The materials for quantitative tasks included both prose and document structures as there are no structures that are unique to quantitative tasks. The majority of these tasks were based on document structures.

Across the entire pool of tasks, the most prevalent structure used for tasks was tables—33 percent of the materials were tables (Table 4-1). While it may seem that there was a disproportionate number of tables, this particular structure comprises a wide range of materials that present information in matrix formats using words, numbers, pictures, and symbols. Thus, materials such as transportation schedules, menus, tables of contents, as well as tables of information, were categorized as tables.

Table 4-1. Percentages of stimulus materials by categories of structures

Structure	Percent of Tasks		Total
	Original in 1985	New in 1992	
Exposition	6	15	21
Narrative and Poetry	1	5	6
Tables	23	10	33
Charts and Graphs	4	6	10
Forms	13	6	19
Maps	1	2	3
Miscellaneous	4	4	8

4.3.3 Adult Contexts/Content

Since adults do not read printed materials in a vacuum, but rather within a particular context or for a particular purpose, materials were used that represent a variety of contexts or content. Six adult context/content areas were identified as follows:

- **Home and family**: interpersonal relationships, personal finance, housing, and insurance;
- **Health and safety**: drugs and alcohol, disease prevention and treatment, safety and accident prevention, first aid, emergencies, and staying healthy;
- **Community and citizenship**: community resources and being informed;
- **Consumer economics**: credit and banking, savings, advertising, making purchases, and maintaining personal possessions;
- **Work**: occupations, finding employment, finance, and being on the job; and
- **Leisure and recreation**: travel, recreational activities, and restaurants.

An attempt was made to include as broad a range of contexts and contents as possible and to select materials that would not be so specialized as to be familiar only to certain groups. This was to ensure that any disadvantages for people with limited background knowledge would be minimized.

Across the entire pool of tasks, 32 percent of the materials fell into the community/citizenship category (Table 4-2). While it may seem that this category is over-represented, it is a very broad category and includes such materials as news articles from newspapers and magazines, information from governmental agencies, transportation schedules, information from schools and colleges, and so on.

Table 4-2. Percentages of tasks by categories of context/content

Context/Content	Percent of Tasks		Total
	Original in 1985	New for 1992	
Home/Family	7	7	14
Health/Safety	3	1	4
Community/Citizenship	12	20	32
Consumer Economics	11	5	16
Work	13	2	15
Leisure/Recreation	6	13	19

The materials and contexts described above define the axes of the matrix in Table 4-3. This table illustrates that the tasks included in the assessment were based on a variety of materials from a variety of contexts. Each dot indicates that at least one task was included that was based on a particular kind of material from a particular context. For example, the row for the content area labeled health/safety contains two dots, one under exposition and one under tables. This means the assessment included tasks that were based on two types of materials, exposition and tables, related to the context of health/safety.

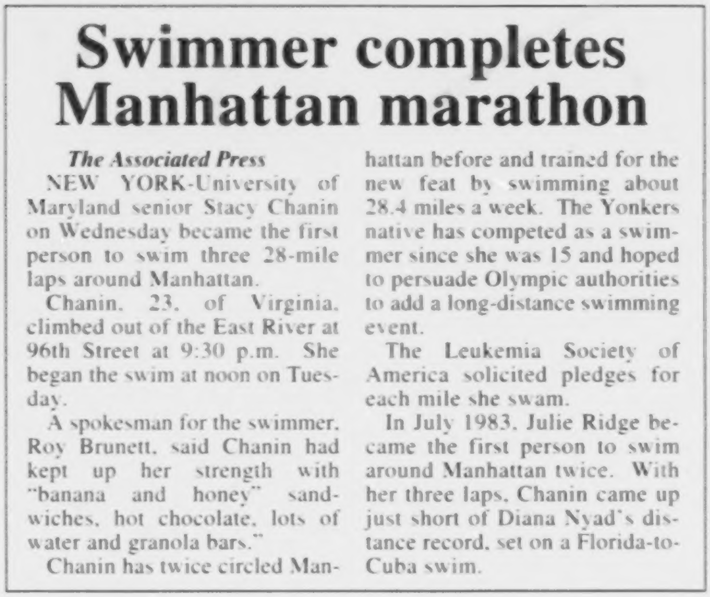
Table 4-3. Task coverage by context or content and type of material

	Materials						
Context/Content	Exposition	Narrative/ Poetry	Tables	Charts/ Graphs	Forms	Maps	Miscellaneous
Home/Family	✓		✓		✓		✓
Health/Safety	✓		✓				
Community/ Citizenship	✓		✓	✓	✓	✓	
Consumer Economics	✓		✓	✓	✓		✓
Work	✓		✓	✓	✓		✓
Leisure/Recreation	✓	✓	✓	✓		✓	✓

4.3.4 Processes/Strategies

After the stimulus materials were selected, tasks were developed that simulated the way people would use the materials and required different strategies for successful task completion. Prose tasks were developed that involve three strategies for processing information: *locating*, *integrating*, and *generating* information. For *locating* tasks, readers must match information given in the question with either literal or synonymous information in the text (see Exhibit 4-1, "swimmer" tasks).

Exhibit 4-1. Example of prose locating task



Swimmer completes Manhattan marathon

The Associated Press

NEW YORK-University of Maryland senior Stacy Chanin on Wednesday became the first person to swim three 28-mile laps around Manhattan.

Chanin, 23, of Virginia, climbed out of the East River at 96th Street at 9:30 p.m. She began the swim at noon on Tuesday.

A spokesman for the swimmer, Roy Brunett, said Chanin had kept up her strength with "banana and honey" sandwiches, hot chocolate, lots of water and granola bars."

Chanin has twice circled Manhattan before and trained for the new feat by swimming about 28.4 miles a week. The Yonkers native has competed as a swimmer since she was 15 and hoped to persuade Olympic authorities to add a long-distance swimming event.

The Leukemia Society of America solicited pledges for each mile she swam.

In July 1983, Julie Ridge became the first person to swim around Manhattan twice. With her three laps, Chanin came up just short of Diana Nyad's distance record, set on a Florida-to-Cuba swim.

Find the article "Swimmer completes Manhattan marathon" on page 2 of the newspaper provided and answer the following questions.

11. Underline the sentence that tells what Ms. Chanin ate during the swim.

12. At what age did Chanin begin swimming competitively?____

Of the original prose tasks, about one-third were *locating* tasks, and of the new prose tasks developed for the survey, about two-thirds were *locating* tasks. Of the total item pool—the original and new combined—slightly over half the tasks require readers to use *locating* strategies.

Integrating tasks require readers to pull together two or more pieces of information located at different points in the text. None of the original prose tasks were *integrating* tasks, and of the new prose tasks developed for the survey, about one-fourth were *integrating* tasks.

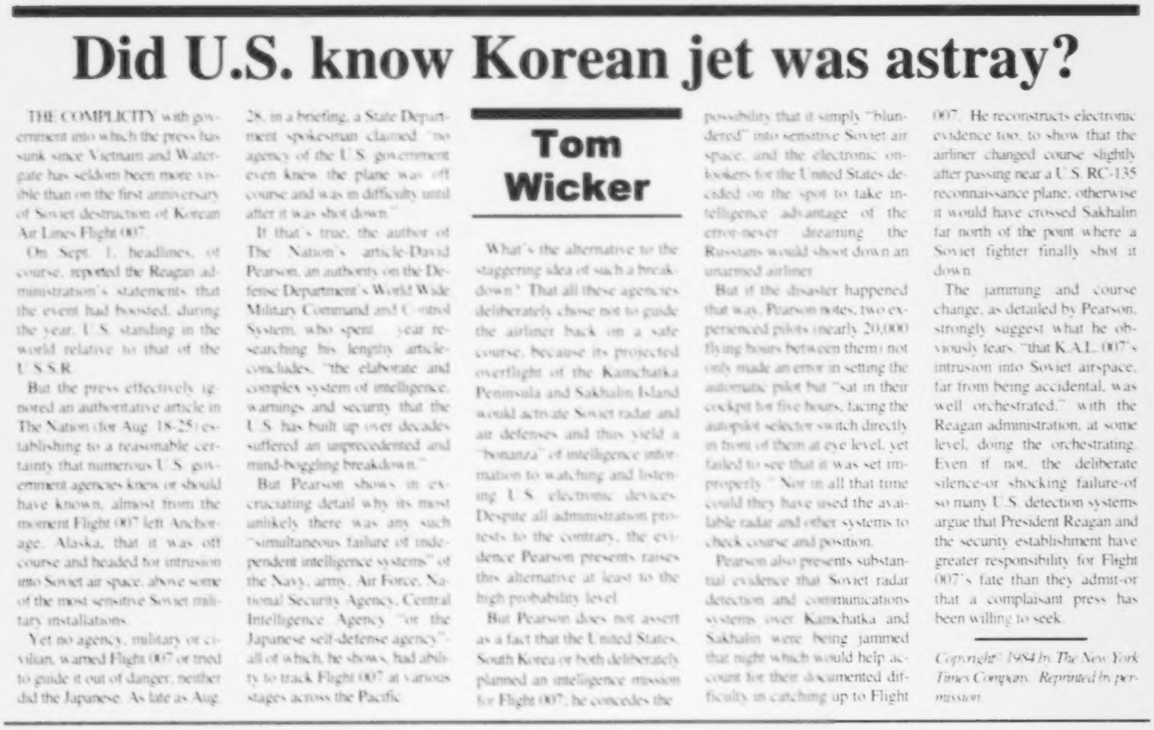
Generating tasks require readers not only to process information located at different points in the text, but also to go beyond that information by making broad, text-based inferences in order to produce new information (see Exhibit 4-2, "Dickinson" task) or by drawing on their knowledge about a subject (see Exhibit 4-3, "Wicker" task). Of the original prose tasks, about two-thirds were *generating* tasks. Of the new prose tasks developed for the survey, about one-tenth were *generating* tasks. Of the total item pool—the original and new combined—just under a third were *generating* tasks.

Exhibit 4-2. Example of prose generating task

The pedigree of honey
Does not concern the Bee—
A clover, any time, to him
Is Aristocracy— (Emily Dickinson)

11. What is the poet trying to express in this poem?________________

Exhibit 4-3. Example of prose generating task (reduced from original size)



Did U.S. know Korean jet was astray?

Tom Wicker

THE COMPLICITY with government into which the press has sunk since Vietnam and Watergate has seldom been more visible than on the first anniversary of Soviet destruction of Korean Air Lines Flight 007.

On Sept. 1, headlines, of course, reported the Reagan administration's statements that the event had boosted, during the year, U.S. standing in the world relative to that of the U.S.S.R.

But the press effectively ignored an authoritative article in The Nation (for Aug. 18-25) establishing to a reasonable certainty that numerous U.S. government agencies knew or should have known, almost from the moment Flight 007 left Anchorage, Alaska, that it was off course and headed for intrusion into Soviet air space, above some of the most sensitive Soviet military installations.

Yet no agency, military or civilian, warned Flight 007 or tried to guide it out of danger; neither did the Japanese. As late as Aug. 28, in a briefing, a State Department spokesman claimed "no agency of the U.S. government even knew the plane was off course and was in difficulty until after it was shot down."

If that's true, the author of The Nation's article-David Pearson, an authority on the Defense Department's World Wide Military Command and Control System, who spent year researching his lengthy article-concludes, "the elaborate and complex system of intelligence, warnings and security that the U.S. has built up over decades suffered an unprecedented and mind-boggling breakdown."

But Pearson shows in excruciating detail why its most unlikely there was any such "simultaneous failure of independent intelligence systems" of the Navy, army, Air Force, National Security Agency, Central Intelligence Agency "or the Japanese self-defense agency"-all of which, he shows, had ability to track Flight 007 at various stages across the Pacific.

What's the alternative to the staggering idea of such a breakdown? That all these agencies deliberately chose not to guide the airliner back on a safe course, because its projected overflight of the Kamchatka Peninsula and Sakhalin Island would activate Soviet radar and air defenses and thus yield a "bonanza" of intelligence information to watching and listening U.S. electronic devices. Despite all administration protests to the contrary, the evidence Pearson presents raises this alternative at least to the high probability level.

But Pearson does not assert as a fact that the United States, South Korea or both deliberately planned an intelligence mission for Flight 007; he concedes the possibility that it simply "blundered" into sensitive Soviet air space, and the electronic onlookers for the United States decided on the spot to take intelligence advantage of the error-never dreaming the Russians would shoot down an unarmed airliner.

But if the disaster happened that way, Pearson notes, two experienced pilots (nearly 20,000 flying hours between them) not only made an error in setting the automatic pilot but "sat in their cockpit for five hours, facing the autopilot selector switch directly in front of them at eye level, yet failed to see that it was set improperly." Nor in all that time could they have used the available radar and other systems to check course and position.

Pearson also presents substantial evidence that Soviet radar detection and communications systems over Kamchatka and Sakhalin were being jammed that night which would help account for their documented difficulty in catching up to Flight 007. He reconstructs electronic evidence too, to show that the airliner changed course slightly after passing near a U.S. RC-135 reconnaissance plane; otherwise it would have crossed Sakhalin far north of the point where a Soviet fighter finally shot it down.

The jamming and course change, as detailed by Pearson, strongly suggest what he obviously fears, "that K.A.L. 007's intrusion into Soviet airspace, far from being accidental, was well orchestrated," with the Reagan administration, at some level, doing the orchestrating. Even if not, the deliberate silence-or shocking failure-of so many U.S. detection systems argue that President Reagan and the security establishment have greater responsibility for Flight 007's fate than they admit-or that a complaisant press has been willing to seek.

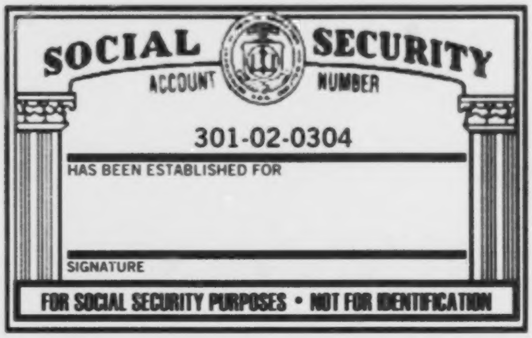
Copyright © 1984 by The New York Times Company. Reprinted by permission.

Find the article "Did U.S. know Korean jet was astray?" on the front page of the newspaper provided and answer the question below.

8. What argument is Tom Wicker making in his column?________________

The strategies required by document tasks also include locating, integrating, and generating information as well as cycling through information. For *locating* tasks, readers must match one feature or category of information given in the task with either identical or synonymous information in a document. (see Exhibit 4-4, "Social Security card" task). About two-thirds of the original document tasks and about two-thirds of the new document tasks were *locating* tasks. Thus, about two-thirds of the total document pool were *locating* tasks.

Exhibit 4-4. Example of document locating task



SOCIAL SECURITY
ACCOUNT NUMBER
301-02-0304
HAS BEEN ESTABLISHED FOR
SIGNATURE
FOR SOCIAL SECURITY PURPOSES • NOT FOR IDENTIFICATION

1. Here is a Social Security card. Sign your name on the line that reads "signature."

[Note: The critical element in scoring this task was not a proper signature, but successfully locating the place where the signature belongs.]

Cycling tasks require the reader to repeat the matching process by identifying all instances that satisfy a set of conditions stipulated in the question or directive (see Exhibit 4-5, "employment form" task). About one-ninth of the original document tasks, but none of the new document tasks were *cycling* tasks. Of the total document literacy pool, about one-tenth were *cycling* tasks.

Exhibit 4-5. Example document cycling task

You have gone to an employment center for help in finding a job. You know that this center handles many different kinds of jobs. Also, several of your friends who have applied here have found jobs that appeal to you.

The agent has taken your name and address and given you the rest of the form to fill out. Complete the form so the employment center can help you get a job.

Birth date______________ Age_____ Sex: Male_____ Female_____

Height____________ Weight____________ Health____________

Last grade completed in school_____________

Kind of work wanted:

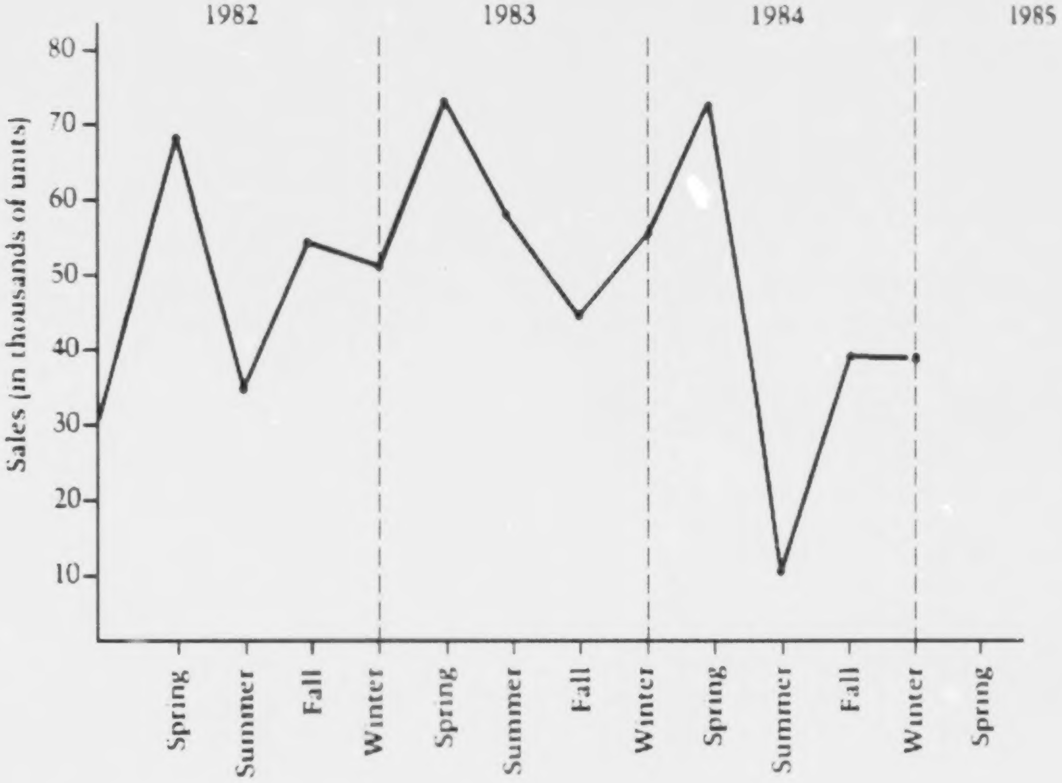
Part-time__________ Summer____________

Full-time__________ Year-round_________

[Note: this document was scored as two tasks: one for entering all personal elements (birth date, age, sex, height, weight, health, and schooling) and another for entering the two features of the kind of work wanted. The later task did not fit the IRT scale and was not included in figuring document literacy scale scores.]

To complete *integrating* tasks, readers must either match on two or more features located in different parts of the document or compare and/or contrast information (see Exhibit 4-6, "graph" task). About one-ninth of the original, and one-fourth of the new document tasks were *integrating* tasks. Of the total document pool, about one-seventh were *integrating* tasks.

Exhibit 4-6. Example document integrating task



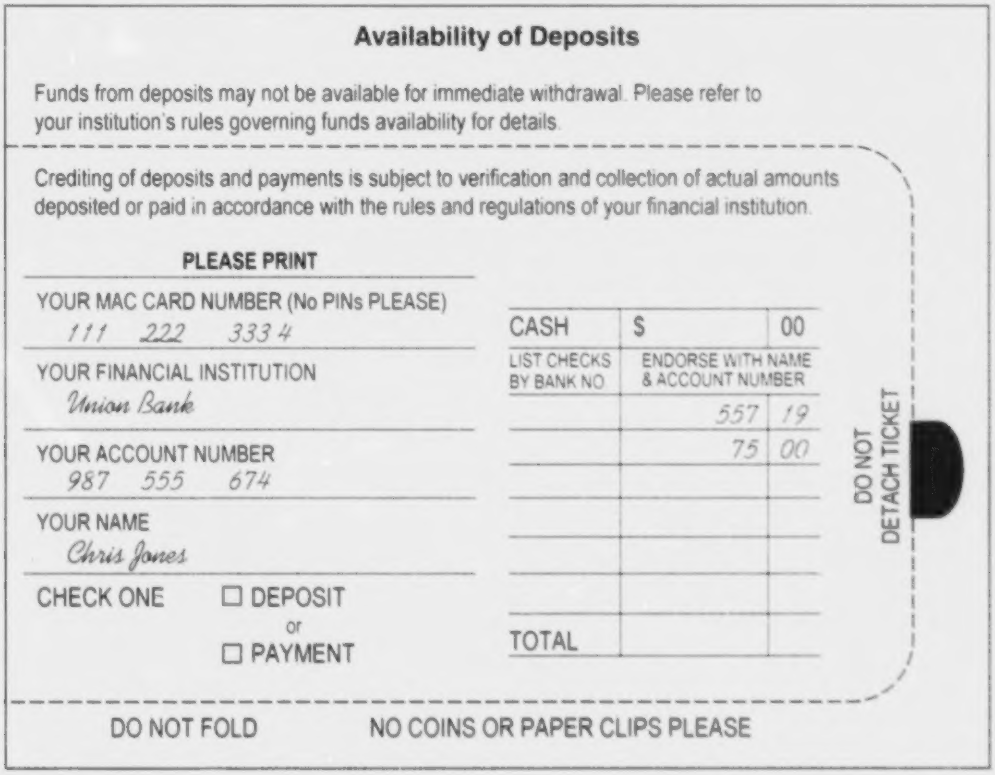
13. You are a marketing manager for a small manufacturing firm. This graph shows your company's sales over the last three years. Given the seasonal pattern shown on the graph, predict the sales for Spring 1985 (in thousands) by putting an "X" on the graph.

As with *generating* tasks in the prose domain, *generating* tasks involving documents require readers to go beyond information in the document either by drawing on their knowledge of the subject or by making inferences to produce new information. About one-ninth of the original, and one-tenth of the new document tasks were *generating* tasks. Of the total document pool, about one-tenth were *generating* tasks.

Quantitative tasks require readers to perform arithmetic operations—addition, subtraction, multiplication, or division—either singly or in combination. Some quantitative tasks require readers to explain how they would solve a problem rather than just to produce a numerical answer, and others require the use of a simple, four-function calculator to solve the problem. Tasks can be more or less difficult for readers depending on the type of arithmetic operation involved, the ease of determining what operations were needed, and the ease of locating or identifying the appropriate numbers. Among the National Adult Literacy Survey tasks, the representation of numerical information associated with the quantitative tasks included whole numbers, decimals, percentages, fractions, and time (hours and minutes).

Addition and subtraction tasks are usually considered the easiest operations (see Exhibit 4-7, "deposit slip" task). Of the original quantitative tasks, about one-fourth each involved the operations of addition and subtraction. Of the new quantitative tasks, about one-fifth were addition and somewhat more than one-fifth were subtraction tasks. Across the total quantitative pool, about one-fourth each were addition and subtraction tasks.

Exhibit 4-7. Example quantitative addition task



Availability of Deposits

Funds from deposits may not be available for immediate withdrawal. Please refer to your institution's rules governing funds availability for details.

Crediting of deposits and payments is subject to verification and collection of actual amounts deposited or paid in accordance with the rules and regulations of your financial institution.

PLEASE PRINT

YOUR MAC CARD NUMBER (No PINs PLEASE)
111 222 333 4

YOUR FINANCIAL INSTITUTION
Union Bank

YOUR ACCOUNT NUMBER
987 555 674

YOUR NAME
Chris Jones

CHECK ONE ☐ DEPOSIT or ☐ PAYMENT

CASH	$	00
LIST CHECKS BY BANK NO	ENDORSE WITH NAME & ACCOUNT NUMBER	
	557	19
	75	00
TOTAL		

DO NOT DETACH TICKET

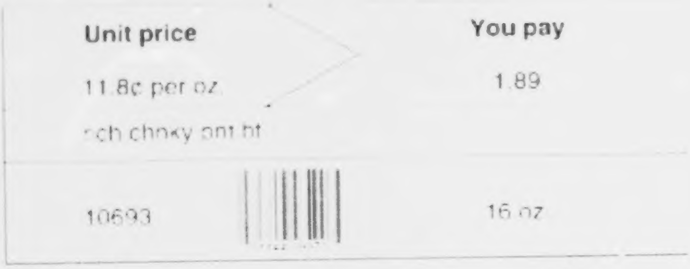
DO NOT FOLD NO COINS OR PAPER CLIPS PLEASE

5. You wish to use the automatic teller machine at your bank to make a deposit. Figure the total amount of the two checks being deposited. Enter the amount on the form in the space next to TOTAL.

Multiplication and division tasks are usually considered more difficult than addition or subtraction tasks (see Exhibit 4-8, "cost per ounce" task). About one-sixth of the original quantitative tasks were evenly divided between the operations of multiplication and division. Of the new quantitative tasks, about one-fifth were multiplication and somewhat fewer than one-fifth were division tasks. Across the total quantitative pool, about one-fourth of the tasks involved the operations of multiplication and division.

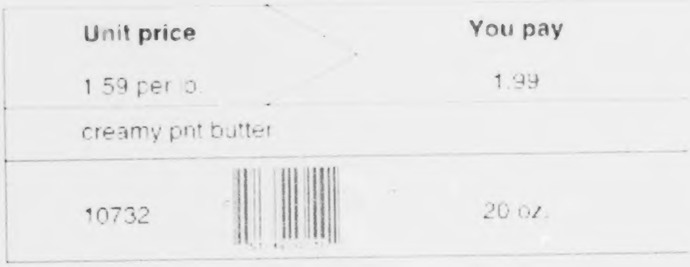
Exhibit 4-8. Example quantitative division task

A.



Unit price	You pay
11.8¢ per oz.	1.89
rch chnky pnt bt	
10693	16 oz

B.



Unit price	You pay
1.59 per lb.	1.99
creamy pnt butter	
10732	20 oz.

You need to buy peanut butter and are deciding between two brands.

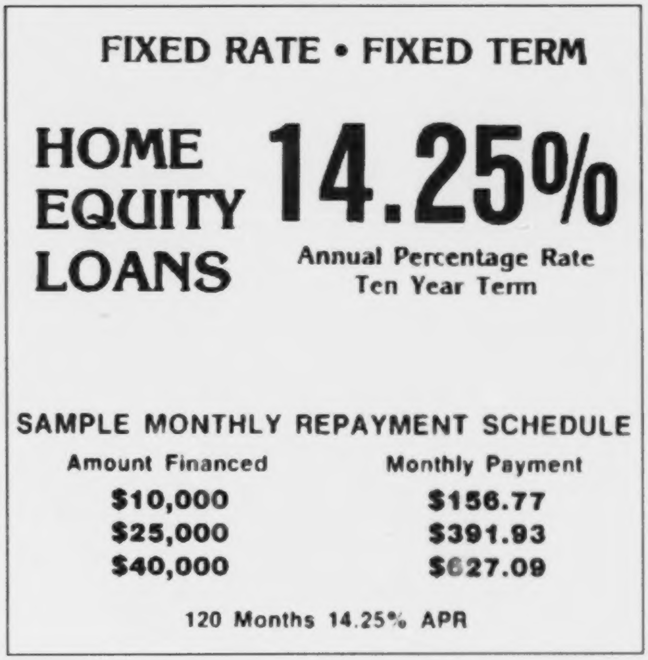
2. Estimate the cost per ounce of the creamy peanut butter. Write your estimate on the line provided.

____

Tasks that require more than one operation are considered even more difficult (see Exhibit 4-9, "home equity loan" task). About one-third of the original and one fifth of the new quantitative tasks involved a combination of operations. Across the total quantitative pool, about one-fourth were combination tasks.

Other factors are also associated with task difficulty. Deciding what operation is appropriate is sometimes obvious from the wording (see Exhibit 4-7, "deposit slip" task) but sometimes indirect, requiring readers to infer which operation they should perform (see Exhibit 4-9, "home equity loan" task).

Exhibit 4-9. Example quantitative combination task



FIXED RATE • FIXED TERM

HOME EQUITY LOANS

14.25%

Annual Percentage Rate
Ten Year Term

SAMPLE MONTHLY REPAYMENT SCHEDULE

Amount Financed	Monthly Payment
$10,000	$156.77
$25,000	$391.93
$40,000	$627.09

120 Months 14.25% APR

3. You need to borrow $10,000. Find the ad for Home Equity Loans on page 2 in the newspaper provided. Explain to the interviewer how you would compute the total amount of interest charges you would pay under this loan plan. Please tell the interviewer when you are ready to begin.

Similarly, sometimes the numbers that are required to perform the operation are easily identified (see Exhibit 4-7, "deposit slip" task), while for other tasks the required numbers to use in setting up the problem may be embedded in text that has distractors—related but incorrect numbers that might confuse the reader (see Exhibit 4-9, "home equity loan" task).

The materials and processes described above for prose, document, and quantitative literacy tasks define the rows and columns in Table 4-4. The cells with a dot indicate that tasks with that particular combination of material and process were included in the pool of literacy tasks for the National Adult Literacy Survey. For example, some tasks based on expository materials required subtraction, but there were no expository-based tasks requiring addition. The design for the survey did not require that tasks cover all possible combinations of materials and processes.

Table 4-4. Task coverage by process and type of material

Process	Materials						
	Exposition	Narrative/ Poetry	Tables	Charts/ Graphs	Forms	Maps	Miscellaneous
Locate	✓	✓	✓	✓	✓	✓	✓
Integrate	✓	✓	✓	✓		✓	
Generate	✓	✓	✓	✓			
Cycle					✓	✓	✓
Add			✓	✓	✓	✓	✓
Subtract	✓		✓	✓	✓		✓
Multiply	✓			✓	✓		
Divide	✓		✓	✓			✓
Combination			✓		✓		✓

Given the strategies required for processing information, the tasks were open-ended rather than multiple choice. That is, they required readers to engage in activities that are similar to those they might perform if they actually encountered the materials and, thus, were not constrained by an artificial set of response requirements. For example, tasks included reading and responding to editorials, news stories, and classified listings in a newspaper; writing a letter to a credit department; explaining the differences between two types of job benefits; completing a bank deposit slip; writing a check; keeping a running balance in a check ledger; and filling out a form to order merchandise from a catalog.

Because the tasks were open-ended, they required a variety of response modes. For some tasks, the respondents were asked to underline or circle information in the stimulus or copy information from it. For tasks that required completing a form, respondents copied information from the directive or question onto the form. In some cases, the information to be copied involved numbers that were then used to perform an arithmetic operation. Other tasks required respondents to produce an answer, such as making inferences based on information in the stimulus or explaining how to set up and solve a quantitative problem. Incorporating a variety of response modes ensured that the simulation tasks reflected real-life uses of printed materials.

4.3.5 Task Difficulty

Each of the types of tasks described above extends over a range of difficulty on the three scales. Research on the Young Adult Literacy Assessment and Workplace Literacy Survey tasks revealed that the difficulty of a particular task is a result of the interaction of the type of process or strategy required by the task with other variables. For the prose and document tasks these other variables include:

- The number of categories or features of information in the directive that the reader has to process;
- The number of categories or features of information in the text or document that can serve as distractors or plausible answers;

- The degree to which the information given in the question has less obvious identity with the information stated in the text or document; and
- The length and density of the text or the structure of the document.

An analysis of quantitative tasks has shown that the information processing required to complete the tasks affects their difficulty. In general, it appears that many adults can perform simple arithmetic operations when both the numbers and the types of operation are made explicit. The tasks become increasingly difficult, however, when these same operations are performed on numbers that must be located and extracted from different types of texts or documents that contain plausible but irrelevant numbers, or when these operations must be inferred from the directive. As a result, the difficulty of quantitative tasks seems to be a function of:

- The particular operation called for;
- The number of operations needed to perform the task;
- The extent to which the numbers are embedded in printed materials; and
- The extent to which an inference must be made to identify the type of operation to perform.

Because this survey was being administered to a nationally representative sample, it was important to capture the full range of literacy skills that people possess and not just to focus on those adults who may have low-level literacy skills. The tasks included in the survey, therefore, covered a range of difficulty across each of the scales. During the development of the new tasks, the variables described above were taken into account to ensure a range of difficulty, thus extending and refining the literacy scales as represented by the tasks from the young adult literacy assessment.

4.3.6 Development of Scoring Guides

As the new tasks were developed, scoring guides were written specifying correct responses to the tasks. Guides for many of the tasks included the following score points:

1: correct answer
2: incorrect answer
9: response of "don't know"
0: no response or blank

Guides for some of the tasks, particularly the *generating* tasks, delineated a finer breakdown of score points. The purpose in doing so was to be able to provide data on various correct and incorrect responses to tasks that might be of interest to researchers. Thus, for example, the scoring guide for the Dickinson poem is as follows:

1: no response written or blank
2: literal interpretation
*3: thematic interpretation
9: response of "don't know"
0: no response or blank
* correct response

As another example, the guide for the home equity loan task is as follows:

1: The respondent states something other than an explanation of computing the interest charges or gives an incorrect explanation
2: The respondent explains one but not both of the steps in computing the total interest charges or is vague about the steps
*3: The respondent explains the two basic steps in computing the total interest charges
9: response of "don't know"
0: No response
* correct response

The scoring guides for the tasks developed for the 1992 assessment underwent several stages of verification and revision. During the test development stage, the tasks underwent a test specialist review, part of which involved checking the accuracy and completeness of the scoring guides. When the scoring was done for the field test of the new tasks, the scoring guides were revised so they would reflect the kinds of responses that people were making to the tasks. As a result of the field test, some of the tasks as well as their scoring guides were revised. In addition, some scoring guides were further revised when the first responses from the main data collection were received. The scoring guides for the tasks from the young adult survey were exactly the same as those used for scoring the tasks for that survey.

4.3.7 Assembling the Tasks for Administration

From a pool of about 110 new tasks developed for the survey, 81 tasks were selected and assembled into seven blocks or sections. Each block was designed to take about 15 minutes of administration time. In selecting the tasks and assembling the new blocks, the following factors were taken into account:

- The inclusion of roughly an equivalent number of tasks from each of the three literacy scales;
- The inclusion of a broad range of content from the identified adult contexts;
- The inclusion of a wide variety of materials and structures;
- A range of difficulty across the tasks as determined from field-test data;
- Representation of content relating to various racial/ethnic groups;
- A variety of response modes; and
- The assignment of all the quantitative tasks requiring the use of a calculator to one block.

Of the new tasks that were selected for the final survey, 27 were selected from the prose scale, 26 from the document, and 28 from the quantitative. These tasks were distributed as evenly as possible across the seven new blocks. Comparatively, the 1985 survey had 14 prose items, 56 document items and 15 quantitative items. Because the new item pool could in and of itself become the basis of a future assessment, it was deemed more important to include a balanced number of new tasks from each scale rather than to achieve balance across the entire pool of both original and new tasks.

A balanced representation of racial/ethnic groups was achieved across the entire set of stimulus materials used in the survey—the ones for the newly developed tasks plus the original materials from the Young Adult Literacy Assessment—not just within one block. About 55 percent of the stimulus materials were neutral with respect to both gender and race/ethnicity—that is, they did not contain any references to people. In the remaining materials, the references to men and women were about equal, and references to specific racial/ethnic minority groups were found in about 25 percent of the materials. In the remaining 75 percent, the references were either neutral with respect to race/ethnicity or the race/ethnicity of the person referred to was identifiable only if someone might have background knowledge about that particular person.

In addition to seven blocks of new tasks, a core set of six literacy tasks—two from each of the three scales—was assembled. These tasks were relatively easy and served to ease transition from background tasks to easier tasks. The core set was designed to take 5 to 10 minutes to complete. The entire survey was designed to take approximately an hour to complete.

The full set of 166 tasks, assembled into 13 blocks and the core, ensured broad, balanced, and representative coverage of materials and content; however, it would take about three and a half hours for each respondent to complete that number of tasks. Because about 45 minutes seemed to be a reasonable amount of time to expect respondents to spend on the literacy tasks, some form of item sampling procedure was essential. The design most suitable for this purpose is a powerful variant of standard matrix sampling called balanced incomplete block (BIB) spiraling. In BIB spiraling, as in standard matrix sampling, no respondent is administered all of the tasks in the assessment pool. Unlike standard matrix sampling, however, in which items or tasks are assembled into discrete booklets, BIB spiraling allows for the estimation of relationships among all the tasks in the pool through the unique linking of blocks.

With the BIB spiral design, the 13 blocks of tasks—the seven new blocks and the six old blocks—were assembled into 26 assessment booklets, each of which contained a unique combination of three blocks. In addition, each booklet included the section of core tasks. The application of the BIB design resulted in the configuration of booklets shown in Table 4-5. In this design, each block appeared with the same frequency—in six of the 26 booklets—and each block was paired one time with every other block. Position effects were also controlled for at the block level since each block appeared twice in each of the possible positions in the booklets—first, middle, and last. On the three National Adult Literacy Survey data files, the booklet number is identified in the variable BOOK, the category labels of which identify the blocks by letter code (rather than number as shown in the following table).

Table 4-5. Balanced Incomplete Block design for 26 booklets

Booklet number	Core	Block numbers contained in booklet		
1	C	1	2	13
2	C	2	3	9
3	C	3	4	7
4	C	4	13	8
5	C	13	9	6
6	C	9	7	10
7	C	7	8	11
8	C	8	6	12
9	C	6	10	5
10	C	10	11	1
11	C	11	12	2
12	C	12	5	3
13	C	5	1	4
14	C	1	3	8
15	C	2	4	6
16	C	3	13	10
17	C	4	9	11
18	C	13	7	12
19	C	9	8	5
20	C	7	6	1
21	C	8	10	2
22	C	6	11	3
23	C	10	12	4
24	C	11	5	13
25	C	12	1	9
26	C	5	2	7

The spiral component of the design ordered the books for administration so that each booklet was completed by a random sample of respondents. Thus, each booklet and each block was completed by approximately the same number of respondents (Table 4-6). One outcome of the BIB spiral design is that every task is taken by a randomly equivalent subsample of respondents. This ensures that reliable estimates of population performance can be calculated for every task. An additional benefit of this methodology is that every pair of tasks is taken by a representative subsample of the total sample so that correlations between pairs of tasks can be estimated.

Table 4-6. Number of persons responding to each booklet and to each block

Booklet	Number	Block	Number
1	1,000	1	5,748
2	963	2	5,792
3	947	3	5,675
4	973	4	5,683
5	964	5	5,558
6	963	6	5,761
7	947	7	5,598
8	963	8	5,765
9	971	9	5,703
10	1,000	10	5,766
11	966	11	5,782
12	893	12	5,598
13	904	13	5,752
14	965		
15	968		
16	953		
17	969		
18	916		
19	933		
20	941		
21	984		
22	954		
23	922		
24	946		
25	938		
26	911		

Chapter 5

THE HOUSEHOLD SURVEY

Susan Rieger and Martha Berlin, Westat, Inc.

5.1 OVERVIEW

To gather information on adults' literacy skills, trained staff interviewed a nationally representative sample of about 13,600 individuals age 16 and older, residing in private households and college dormitories across the United States. Survey participants had been randomly selected to represent the adult population in the country as a whole. Individuals from black and Hispanic groups were sampled at about double the rate of the remainder of the population to assure reliable estimates of their literacy proficiencies. In addition, the sample contained enough individuals in the 21 to 25 age range to allow comparisons with the 1985 Young Adult Literacy Assessment. (The 1985 study was conducted by Educational Testing Service under a grant from the National Center for Education Statistics as an adjunct to the National Assessment of Educational Progress.)

To give states an opportunity to explore the skill levels of their populations, each of the 50 states was invited to participate in a concurrent assessment. While many states expressed an interest, eleven elected to participate in the State Adult Literacy Survey. Approximately 1,000 adults ages 16-64 were surveyed in each of the following states: California, Illinois, Indiana, Iowa, Louisiana, New Jersey, New York, Ohio, Pennsylvania, Texas, and Washington. (Florida also participated in the state survey, but its data collection was unavoidably delayed until 1993 and is not reflected in this report.) To permit comparisons of the state and national results, the survey instruments administered to the state and national samples were identical, and the data were gathered at the same time.

Finally, more than 1,100 inmates in 87 federal and state prisons were included in the survey. Their participation helped to provide better estimates of the literacy levels of the total population and make it possible to report on the literacy proficiencies of this important segment of society. To ensure comparability with the national survey, the simulation tasks given to the prison participants were the same as those given to the household survey population. However, to address issues of particular relevance to the prison population, a revised version of the background questionnaire was developed.

The literacy survey was conducted from February through August, 1992. More than 400 trained interviewers, some of whom were bilingual in English and Spanish, visited households to select and interview adults. Each survey participant was asked to spend approximately one hour responding to a series of diverse literacy tasks as well as questions about his or her demographic characteristics, educational background, reading practices, and other areas related to literacy. Based on their responses to the survey tasks, adults received proficiency scores along three scales that reflect varying degrees of skill in prose, document, and quantitative literacy.

Following the completion of an interview, interviewers edited all materials for legibility and completeness. The interviewers sent their completed work to their regional supervisors for a complete edit of the instruments, quality control procedures, and any required data retrieval. As these tasks were completed, the cases were shipped to ETS for processing.

In accordance with the contract between the ETS and the National Center for Education Statistics, a field test was conducted in the spring of 1991 using a sample of some 2,000 adults drawn from 16 primary sampling units (PSUs). The purposes of the field test were as follows:

- To evaluate the impact of incentives on response rates, performance, and survey costs;
- To evaluate newly developed literacy exercises for item bias and testing time; and
- To evaluate the administration and appropriateness of the background questions.

5.2 LISTING

The implementation of an area probability design such as the one used for the National Adult Literacy Survey requires the development of a list of dwelling units in each second-stage sampling unit, or area segment. This section describes the procedures used to carry out the address listing operation. For the national sample, the National Adult Literacy Survey design involved 101 primary sampling units (PSUs) and 2,064 area segments, while the design for the state samples involved 109 PSUs and 1,837 area segments. Hence, the total household sample was distributed across 210 PSUs and 3,901 area segments.[1]

The survey listing operation was precedent setting in at least two aspects. First, the project was one of the first to use the Census Bureau's Topologically Integrated Geographical Encoding and Referencing (TIGER) System file to computer generate segment maps for use in the listing phase of an area probability sample design. In addition, the sheer size of the operation made it unique—it was arguably one of the largest listing efforts ever carried out for a single national survey. During an 11-week period in the Fall of 1991, nearly 400,000 dwelling units were listed in the 3,901 area segments. A field organization of nearly 300 people was assembled to carry out the listing operation.

5.2.1 Staff Organization for Listing

The staff for the National Adult Literacy Survey listing operation included 16 supervisors and 272 listers. The supervisors reported to a field manager located in Westat's home office. On average, each supervisor had responsibility for recruiting and supervising 17 listers located in 13 PSUs. Listers were recruited in August 1991. Many of the listers had previously worked for Westat or for other well-known survey

---

[1] Forty-one PSUs in the national and the state samples overlapped. Hence, the sample comprised 169 unique geographic locations.

research organizations. Some listers had worked on the decennial census in their local areas and had fairly recent experience working in the sample segment areas.

5.2.2 Training Listers

Of the 272 listers recruited, more than 40 percent (115 listers) had had listing experience within the past three years. These persons were trained by home study, using a listing manual and home study guide. This document was a reference manual with practice exercises interspersed throughout and a final examination for review at the end. The home-study-only trainees were instructed to read the manual and to complete the exercises and the final examination. They then removed the completed exercises and examination from the guide and mailed them to the appropriate supervisor for review, keeping the remainder of the text for use as a reference manual during the listing operation.

Inexperienced listers attended a one-day, in-person training session that included lectures, an audiovisual training presentation, and field practice. Before attending, these trainees also completed the home study package and brought the exercises and final examination for review and evaluation by training staff. The training session covered fundamental concepts and basic procedures of listing, problematic aspects of listing, special procedures for working in rural areas, and administrative procedures. Listing procedures unique to the National Adult Literacy Survey project were also presented, including instructions for listing group quarters and use of the computer-generated tract maps and segment maps.

In each city in which lister training was held, a segment near the training site had been selected for practice listing. The training staff prepared practice segment folders containing copies of the tract and segment maps, listing sheets, and listing route forms. Before the training session, the staff performed address listing in the practice segment, noting any problem areas. During training, each trainee was required to list the practice segment individually. After everyone had completed the segment, training staff walked through the segment with the trainees, checking the recording and discussing any problems.

5.2.3 Listing Materials

A segment folder was prepared for each sampled segment. The folder contained (1) a tract map; (2) a segment map; (3) listing sheets, on which the lister recorded each address, one to a line; (4) listing route forms; and (5) a form containing general comments and any special instructions.

The tract and segment maps included in each folder were used to define and describe the sample segments, permitting the listing staff to identify on the ground the exact boundaries of the sampled areas. The 11" x 17" tract map provided an overall picture of the location of the segment within a larger geographic area and within the county, with the actual segment boundaries highlighted in yellow; the purpose of this map was to provide the listers with a geographic context to help them locate the segment. The 8.5" x 11" segment map was a more detailed picture, showing all streets and other features of the area to be listed (i.e., the highlighted area on the tract map); this map could be used by listers to sketch diagrams of the dwelling units within the segment.

Tract and segment maps were previously produced manually, by clerks who outlined the sample area boundaries on maps purchased from the Census Bureau. As noted earlier, the National Adult Literacy Survey was among the first large-scale applications of the Census Bureau's TIGER System file for the production of segment maps in connection with an area probability survey. The TIGER file digitizes all intersections of geographic boundaries used in the 1990 census, and this information can be used to computer generate maps of selected blocks, combinations of blocks, or other geographic units. This method of map production cost considerably less, was more accurate, and was much faster to implement than the traditional method. Westat also made some modifications to the map design (e.g., uniform map sizes, additional detail to improve legibility, automatic insertion of convenient map numbers) to enhance the usefulness of the maps for the listing operation. A complete description of the map production procedures is provided in section 2.2.2.4.

5.2.4 The Listing Operation

The listing operation began immediately after training and was completed by mid-December 1991. During the initial stages of listing, the lister located the assigned segment, using the maps from the segment folder and, when necessary, a local map to verify boundaries. Before beginning to record addresses, the lister "cruised" the segment, making an approximate count of the dwelling units in the segment and correcting the segment and tract maps, if necessary.

To keep the listing costs within reasonable bounds, very large area segments—those containing 500 or more dwelling units—were subdivided into smaller areas, or chunks, according to instructions established by the statistical design staff. Then, one of the chunks was selected with probability proportionate to size, as the area to be completely listed. (More detail on chunking procedures has been included in Chapter 2 of this report.) Approximately 220 segments had dwelling unit counts in excess of 500 and were subdivided in this fashion.

If no major problems were encountered while cruising, the lister began the actual address listing operation. The lister started listing in the northwest corner of the segment. The starting point and the direction of travel were indicated on the segment map. As the lister traveled through the segment, following the specified listing route, he or she recorded the address of each dwelling unit on the listing sheet. If no house and/or apartment number was evident for a dwelling unit, the lister recorded a detailed description of the unit and its location. The lister also completed a listing route form, on which he or she recorded each street name in the order traveled and the beginning and ending intersections for each street.

This procedure made it easier to review and check the listing for completeness and facilitated the subsequent use of the listing by interviewers or other members of the field staff.

Because the design of the National Adult Literacy Survey was based on the 1990 decennial census data and the listing operation was carried out less than 2 years after the field operations for the census, relatively few structural changes had occurred in the segments. Hence, in most segments, the difference between the expected and actual numbers of dwelling units was not great. Also, for the most part, segment boundaries were still intact and could be easily located from census maps.

Completed segment listings were returned to the Westat listing department, where they were reviewed for completeness, accuracy, legibility, and adherence to procedures. The address information was then coded, keyed, and entered into the survey control file. Of the 3,901 segments originally selected for listing, eight segments yielded no dwelling units. In the remaining 3,893 segments, nearly 400,000 dwelling units were listed. Of these, approximately 45,000 dwelling units were selected to form the national sample (approximately 25,000 dwelling units) and the state samples (approximately 20,000 dwelling units across the 11 participating states).

5.2.5 Quality Control Procedures

Quality control checks applied to the listing operation included a review of each lister's initial assignment and the implementation of procedures to identify and sample dwelling units and structures missed during listing.

5.2.5.1 Quality control of listing sheets

Each lister was required to mail his or her first two completed segment listings to the supervisor for review before working on additional segments. The supervisor reviewed the listings for completeness, accuracy, legibility, and adherence to procedures and provided feedback to the lister immediately. Based on this review of a lister's first assignment, the supervisor could then decide on the type and number of segments to assign to that lister.

Section 2.2.3.4 provides the rationale for the missed dwelling unit and missed-structure procedures and describes the way in which segments and structures were selected for these procedures. The procedures were carried out during the interviewing operation and are described in Section 5.2.5.2 below.

5.2.5.2 Quality control of the listing operation

As a check on the completeness of the address listing operation, the survey interviewer performed two procedures to detect and measure omissions in listing and to correct for them and provide an unbiased sample of missed units at the same rate as the rest of the sample. The missed dwelling unit and missed structure procedures were performed during the data collection effort. As the names imply, the procedures separate the detection of missed dwelling units into two parts—identifying missed units within multi-unit structures and detecting completely missed structures. Each of these two procedures is described below.

Missed Dwelling Unit Procedure The missed dwelling unit procedure was conducted in a sample of structures selected during the listing operation in the home office. It was designed to detect individual units within those structures that were listed. Such units might be in multi-unit structures, such as apartment buildings or duplexes, or they might be separate dwelling quarters within what appeared to the lister to be a single-family structure.

An instruction in the screener directed the interviewer to check the assignment label to see if the missed dwelling unit procedure was required. If the procedure was required and the dwelling unit had been listed as a single-family unit, the interviewer asked the respondent if there were any other living quarters at the address, such as a basement or attic apartment. In multi-unit structures, the interviewer compared mailboxes and doorbells against the listing sheet and looked around the outside of the structure for additional units or entrances, being particularly careful to look for basement, unnumbered, or out-of-the-way apartments that would be easy to miss.

If no missed units were discovered, the interviewer simply checked a circle on the missed dwelling unit form in the screener to verify that the procedure had been carried out. If 10 or fewer additional units were discovered in any segment, they were automatically added to the sample, and the interviewer began efforts to conduct interviews in the households. If more than 10 missed dwelling units were discovered in any unit, a sub-sampling procedure was used to control the number of additional units added to the sample. When this situation arose, the interviewer called the supervisor for subsampling instructions.

Missed Structure Procedure The missed structure procedure was conducted in a sample of segments. A message on the segment folder instructed the interviewer to perform the procedure. Using the tract and segment maps, listing sheets, and the listing route form, the interviewer recanvassed the entire segment to look for single-family houses or multi-unit structures that were omitted from the listing sheets.

If no missed structures were discovered, the interviewer simply checked a box on the missed structure form to verify that the procedure had been performed. If one or more missed structures were found, the interviewer listed all of the dwelling units they contained on the listing sheet and the missed structure form. If 10 or fewer missed dwelling units were discovered in a structure, they were automatically added to the sample and the interviewer began efforts to interview in the households. If more than 10 missed dwelling units were discovered in a structure, a sub-sampling procedure was used to control the number of units added to the sample. When this situation arose, the interviewer called the supervisor for sub-sampling instructions.

5.3 DATA COLLECTION INSTRUMENTS AND INTERVIEWER MATERIALS

The background questionnaire and the exercise booklet, the primary data collection instruments used in the National Adult Literacy Survey, are described in Chapter 4. What follows is a discussion of the other materials used during data collection, including the screener, the interview guides for the exercise booklets, the non-interview report forms, the interviewer manuals, and various field aids.

5.3.1 The Screener

The screener was used to collect household information and to select one or more members of the household for participation in the background interview and the literacy assessment. The interviewer was given a screener for each sampled dwelling unit in an assigned segment. A computer-generated label on the cover specified the case identification number, the address of the dwelling unit, the line number associated with the dwelling unit on the listing sheet, and the version number of the exercise booklet to be administered. When appropriate, messages on the label instructed the interviewer to conduct the missed dwelling unit and/or missed structure procedures (see section 5.2.5.2) or to implement procedures for oversampling black and Hispanic persons in high-minority segments (i.e., those in which black and/or Hispanic persons accounted for 25 percent or more of the total population; see section 2.2.4). In the latter case, the interviewer was instructed to select household members for participation only if the person listed as the owner or renter of the dwelling unit was black or Hispanic.

Two versions of the screener were utilized for the National Adult Literacy Survey, to reflect the different selection criteria used in the national and state samples. The screener began with a household enumeration that was identical in both versions (Exhibit 5-1). The name of the person (or one of the persons) who owned or rented the dwelling unit was recorded on the first line of the enumeration table.

Then each household member's name, relationship to the homeowner/renter, sex, age, and race/ethnicity were recorded. The questionnaire could be administered to any household member age 16 or older.

The screener also contained procedures for selecting the appropriate respondent(s) in each household. The interviewer had no discretion as to whom to include in the sample. Rather, sampling messages specified which one respondent to select in households having one to three eligible members and which two respondents to select in households having four or more eligible members. The national and state sampling rules reflected the different age eligibilities associated with these two sample designs (age 16 or older for the national sample and 16-64 for the state sample). Because the state sample design did not require the oversampling of racial/ethnic groups, the labels on the screeners used in the state surveys did not contain the oversampling message.

116

Exhibit 5-1. Screener form for household enumeration

HOUSEHOLD COMPOSITION

S-1. Including yourself, how many people live in this household?

______ NUMBER

S-2. What is the name of the person, or one of the persons, who owns or rents this home? (ENTER NAME ON LINE 01 OF ENUMERATION TABLE BELOW.)

BOX 1.	• IF ONLY ONE PERSON LIVES IN THE HOUSEHOLD, GO TO S-5. • OTHERWISE, CONTINUE.

S-3. And the other members of this household – what are their names? Let's begin with everyone related to (PERSON 01). (BE SURE RESPONDENT INCLUDES SELF. ENTER NAME(S) IN TABLE).

S-4. Are there any other people living there who are not related to (PERSON 01)?

YES .... 1 (ENTER NAME(S) IN TABLE; THEN S-5.)
NO 2 (S-5)

S-5. [I have listed (READ NAMES IN ORDER).] Is there anyone else living here now, such as friends, relatives or roomers?

YES .... 1 (ENTER NAME(S) IN TABLE; THEN S-6.)
NO 2 (S-6)

S-6. Are there any college students or children in boarding school who usually live here but who are now living away from home in dormitories, or fraternity or sorority houses?

YES .... 1 (ENTER NAME(S) IN TABLE; THEN S-7.)
NO 2 (S-7)

S-7. Have we missed any other household members now away from home who usually live here, for example, someone away on vacation or business, or in a hospital?

YES .... 1 (ENTER NAME(S) IN TABLE; THEN S-8.)
NO 2 (S-8)

ENUMERATION TABLE AFTER LISTING HOUSEHOLD MEMBERS, ASK S-8 THROUGH S-11 FOR EACH PERSON		S-8. What is (PERSON)'s relationship to (PERSON 01)?	S-9. CODE SEX. (ASK IF NOT OBVIOUS.) Is (PERSON) male or female?		S-10. How old was (PERSON) on (his/her) last birthday?	S-11. HAND CARD A. Which of the groups on this card best describes (PERSON)'s racial and ethnic background? (ENTER LETTER.)	
PERSON #	FIRST NAME LAST NAME		M	F			
01		HOMEOWNER/ RENTER	1	2		LETTER	OTHER (SPECIFY)
02			1	2		LETTER	OTHER (SPECIFY)
03			1	2		LETTER	OTHER (SPECIFY)
04			1	2		LETTER	OTHER (SPECIFY)
05			1	2		LETTER	OTHER (SPECIFY)
06			1	2		LETTER	OTHER (SPECIFY)
07			1	2		LETTER	OTHER (SPECIFY)
08			1	2		LETTER	OTHER (SPECIFY)
09			1	2		LETTER	OTHER (SPECIFY)
10			1	2		LETTER	OTHER (SPECIFY)

(OFFICE USE ONLY

A. WHITE, NOT HISPANIC
B. WHITE, HISPANIC
C. BLACK, NOT HISPANIC
D. BLACK, HISPANIC
E. AMERICAN INDIAN
F. ALASKAN NATIVE
G. PACIFIC ISLANDER
H. ASIAN
I. OTHER (SPECIFY)

To facilitate the validation of screeners and the subsequent follow-up of the case, the interviewer asked the respondent for a telephone number where the respondent could be reached. Finally, the interviewer checked the label to determine whether the missed dwelling unit procedure was required (see section 5.2.5.2); if so, the interviewer followed the missed dwelling unit procedures provided at the end of the screener. The screener also contained a non-interview report form, which is described in section 5.3.3.

A Spanish version of the screener was administered in households in which the household members spoke only Spanish. The average administration time for the screener was 8 minutes.

5.3.2 Interview Guides for Exercise Booklets

Each of the 26 versions of the exercise booklet had a corresponding interview guide, which contained specific instructions for directing the exercise booklet and cued the interviewer when to read an instruction or passage or hand materials to the respondent. The instructions in the guides were formatted in different type fonts and contained icons to present a graphic illustration to the interviewer of how to proceed. Interviewer instructions were printed entirely in capital letters, whereas introductions to be read to the respondent were in italic type. Boxed statements summarized the flow of each task within an exercise block to help the interviewer keep track of the flow of the exercise. Icons were used to indicate various materials to be handed to the respondent, such as a newspaper or calculator, and to instruct the interviewer when to tape record a section of the assessment. In the instructions for the last set of tasks in a block, the respondent was asked to tell the interviewer when he or she had finished the block.

The interview guide contained an observation page on which the interviewer recorded answers to several questions about the respondent and the flow of the exercise block. A sample page from an interview guide is shown as Exhibit 5-2. (See Appendix F for full Interview Observation Guide.)

5.3.3 Non-interview Report Forms

When a sampled respondent did not complete either the screener, the background questionnaire, or the exercise booklet, the interviewer was required to complete a non-interview report form. The information collected in these forms served two important purposes: (1) Field supervisors reviewed the forms to determine the case's potential for conversion and (2) the data collected on the form were processed for non-response analysis.

Exhibit 5-2. Example page of interviewer instructions

Block 9C

INSTRUCTIONS FOR YOU: BEFORE R BEGINS BLOCK 9C, READ THE FOLLOWING DIRECTIONS TO THE RESPONDENT.

READ TO R: *This section contains 10 tasks for you to complete. Follow the written directions for each set of tasks. Some of the tasks, however, require the use of a calculator. Before you begin, I would like to show you how to use this calculator.*

HAND R CALCULATOR AND READ: *Turn on the calculator by pushing the ON/C button. You should see 0 in the window. To enter one dollar twenty-five cents, push 1 point (.) 2 5. You should see 1 point 25. Push ON/C to clear. You should see 0. The ON/C button clears or erases the numbers you have put in the calculator.*
To add 5 plus 8, push 5 plus sign (+) 8 equal sign (=). You should see 13. Push ON/C to clear.
To compute 42 minus 12, push 4 2 minus sign (-) 1 2 equal sign (=). You should see 30. Push ON/C to clear.
To multiply 4 times 8, you should push 4 times sign (X) 8 equal sign (=). You should see 32. Push ON/C to clear.
To divide 72 by 9, push 7 2 division sign (÷) 9 equal sign (=). You should see 8.
Always remember to push clear before you start each problem. Please use the calculator for tasks 1 and 2.

R SHOULD DO TASKS 1 AND 2: OIL BILL, USING THE CALCULATOR.

INSTRUCTIONS FOR YOU: IF R DOES NOT WANT TO USE CALCULATOR, SAY: *We would like you to use the calculator for these questions because one purpose is to see how people do arithmetic when they have a calculator. If you do not want to use the calculator, you may go on to the next set of tasks.* IF R DOES NOT USE THE CALCULATOR FOR TASKS 1 AND 2, PLEASE CHECK BOX. ☐

The screener non-interview report form was completed if the sampled address was vacant or not a dwelling unit or if the interviewer was unable to complete a screener at that address. In the latter case, the interviewer provided information about attempts to contact the household and the reason for non-completion. If the reason was that a household member had refused, the interviewer described the reasons why, in the respondent's own words. The interviewer also provided any other information that might help another interviewer to contact the household and/or to complete the screener.

Similar information was collected for the background questionnaire and exercise booklet. On the first page of the non-interview report form, the interviewer provided the specific reasons why the respondent did not complete that instrument. Some reasons for non-response to the background questionnaire or the exercise booklet were directly related to the issue under study, that is, to the respondent's literacy skills. For this latter reason, interviewers were trained to be very conservative in making assumptions about a respondent's ability to complete the background questionnaire or the exercise booklet; they were instructed to allow the respondent to attempt the questionnaire and the exercises if the respondent was willing.

During training, interviewers were carefully instructed in the meaning of the result codes used on the non-interview report forms. The interviewer's manual also contained material to assist interviewers in making distinctions between the codes. In addition, to discourage interviewers from making assumptions about a respondent's capacity to complete a survey instrument, the interviewer instructions emphasized that disabilities that would prevent the respondent from completing one instrument might not prevent the completion of the other. For example, although a respondent who was unable to hold a conversation in English would be prevented from responding to the background questionnaire, he or she might have sufficient English language reading skills to attempt the exercise booklet. Conversely, although a blind respondent would be unable to complete the exercise booklet, he or she would not be prevented from responding to the background questionnaire if the interviewer read aloud the information on the hand cards used during the interview.

Interviewers were also trained to make careful distinctions between the categories of refusal or breakoff, on the one hand, and mental or physical disability, on the other. If the respondent completed part of the background questionnaire or the exercise booklet but refused to continue because the respondent thought the activity was a waste of time, the interviewer was to assign the code of "partial complete, refusal." However, if a hearing problem prevented the respondent from understanding the questions in the background interview, or if a vision problem interfered with the respondent's ability to read the exercise booklet, the interviewer was to assign the code of "partial complete, physical/mental disability" and describe the situation in the background questionnaire/exercise booklet non-interview report form.

Interviewers were also trained to distinguish between respondents with insufficient reading or writing skills to complete the exercise booklet (whose non-response was to be coded as a "reading/writing barrier") and those who could not complete the exercises because of a physical or mental impairment (to be assigned the "physical/mental disability" code).

5.3.4 Interviewer Manuals

Survey interviewers received a manual of general interviewing techniques and a study-specific manual. Westat's general interviewing techniques manual was mailed to all trainees new to Westat before the start of the training session. The manual is a programmed learning text, with exercises interspersed throughout and a final review examination to be completed by the trainee. Trainees new to Westat are instructed to read the material, complete the practice exercises and final examination, and bring them to an in-person general interviewing techniques session held the day before study-specific training.

The study-specific interviewer's manual was assembled as a loose-leaf notebook, so that pages could be changed or added if necessary. It included an introduction to the survey and an overview of interviewer responsibilities. The text covered field materials and procedures for locating sampled households, contacting respondents, and administering the screener, the background questionnaire, and the exercise booklet. It included the question-by-question specifications for the background questionnaire and the non-interview report forms and a copy of the interview guides for the 26 versions of the exercise booklet. The interviewer's manual also contained information on quality control procedures, record keeping, completing the time and expense report, shipping cases to the supervisor, and reporting to the supervisor. A detailed table of contents and section markers assisted the trainee in locating specific information in the manual.

5.3.5 Field Aids

Survey interviewers utilized field aids for (1) locating and contacting respondents, (2) obtaining respondent cooperation, and (3) conducting the interviews.

5.3.5.1 Aids used for locating and contacting respondents

During the initial phase of interviewing, assignments were made by segment. For each assignment, the interviewer was given a segment folder containing the following materials:

- A computer-generated log of identifying information for all sample dwelling units in the segment, which the interviewer used to record the status of the assignment;
- Computer-generated tract and segment maps;
- The list of dwelling units developed during the listing phase and the listing route form; and
- The missed structure procedure form, if the segment had been selected for this procedure.

For each case in the segment, the interviewer was given a call record folder, to record the result of each contact attempt. Space was provided for interviewer notes about the case, such as the respondent's name and telephone number or directions to the respondent's home. Assignment materials for each case, including the labeled screener, were enclosed in the folder.

Interviewers carried copies of a "Sorry I Missed You" card, which briefly described the study and said that the interviewer would return within the next few days. The card also provided a Westat toll-free number that the respondent could call for more information, and it had space for the interviewer's name and telephone number, if the interviewer wished to provide them.

5.3.5.2 Aids used for obtaining respondent cooperation

An advance letter from the U.S. Department of Education's National Center for Education Statistics and a study brochure were mailed to each sampled dwelling unit shortly before the segment was assigned to an interviewer. The purpose of the letter was to introduce the study in a brief and general way and to notify the residents of the household that a member of Westat's interviewing team would soon call on them. Interviewers were given extra copies of the letter—in English and Spanish—to show to respondents or other household members if they did not recall receiving one in the mail or if they could not read the version sent in the mail.

The study brochure included an endorsement letter from then first lady Barbara Bush and photographs of adults engaged in a variety of reading and writing activities, added more details about the study, and answered some of the most frequently asked questions about the survey. Interviewers were also provided with extra copies of the brochure to give to respondents or other household members if they did not recall receiving one.

Each interviewer wore a photo-identification badge with the interviewer's name, the name of the study, Westat's name and address, and the name of the sponsoring agency. Interviewers also carried a letter of introduction printed on the National Center for Education Statistics' letterhead and were instructed to show the letter to respondents or others in the community if there was any question of the legitimacy of the study.

Three non-response letters were used to persuade reluctant respondents to participate: One letter was intended to show, from a variety of angles, how the study would benefit the general public and the federal government; a second letter addressed the special concerns of older respondents; and a third, brief letter was geared toward persons who might be motivated by a simpler format and approach. Field supervisors reviewed all non-response cases to determine if the case held potential for conversion. If so, the supervisor chose the most appropriate refusal letter and mailed it before the next interviewer contact.

5.3.5.3 Aids used during the interview

To assist the interviewer in administering the background questionnaire and the exercise booklet, several aids were provided for use during the interview. The interviewers received a set of hand cards that contained answer categories for particular questions. At various times during the administration of the background questionnaire, the respondent referred to the cards to answer particular questions. The hand cards were printed in English on one side and in Spanish on the other.

To complete many of the exercise booklet tasks, the respondent was required to use one of several stimulus materials. Each interviewer received an almanac, a hand-held calculator, and a tape recorder, which were given to the respondent as needed to complete the tasks. A mock newspaper was provided with each exercise booklet and interview guide set that required the use of one.

5.4 FIELD ORGANIZATION AND TRAINING

5.4.1 Field Organization

The national, state, and prison survey components were carried out simultaneously by one of the largest field organizations ever assembled for a sample survey effort. The field organization was headed by the National Adult Literacy Survey field director, who reported directly to the Westat project director and who was supported by four home-office field managers and 24 field supervisors located across the United States. The supervisors oversaw an interviewing staff of more than 400 interviewers. This section presents a general description of the field organization and the responsibilities of the staff at each level.

5.4.1.1 Lines of responsibility

The home office staff involved in overseeing the National Adult Literacy Survey field organization included the Westat project director, the field director, and four field managers. The survey field director coordinated all activities in the home office related to field operations and kept in close touch with the four field managers and the ETS staff in dealing with issues of production, cost, response rates, shipment of closed-out work, and other issues.

Each of the field managers had responsibility for five to seven regions. In addition to this oversight responsibility, each was charged with a specific field-related task: (1) distribution of supplies and materials to the field, (2) editing and quality control tasks, (3) automated survey control system (ASCS) maintenance and problem solving, and (4) managing the data collection effort for the prison sample.

For purposes of field operations, the 210 PSUs had been divided into 24 regions, each headed by a regional supervisor who lived in the region. The field supervisor's primary responsibility was overseeing the work of an average of 18 interviewers in his or her region. The supervisor's responsibility also included participation in recruiting and training interviewers for the region. A total of 421 interviewers were recruited in early February 1992. To compensate for attrition and slow production in a small number of areas, an additional 35 interviewers were recruited in April 1992.

Each supervisor was supported in the field by an editor, who was responsible for completely editing each case received from the field and completing a transcription sheet, from screener data, of the characteristics of each sampled unit. (See section 5.6 for a more complete description of the editing tasks.) In addition, the editor assisted the supervisor in such activities as shipping completed work to ETS on a weekly basis and mailing interviewer assignments and materials.

5.4.1.2 Interviewer recruitment

Field staff were recruited and hired directly, and not through interviewing services. Interviewers were hired from the areas in which the interviewing assignments were located. The primary source of potential field staff was Westat's computerized field personnel file containing information on approximately 4,000 persons who have worked on Westat field studies in the previous three years. This flexible system can quickly produce lists by geographic area of available field personnel who meet the qualifications for a project. The system contains demographic information on race, languages spoken, special field skills, and time and geographic availability. Project evaluations are also included in the system, including productivity, accuracy, cooperation, dependability, and length of service for each project. In addition to the computerized file, manual files contain additional information, which was reviewed before recruiting.

When recruiting interviewers, supervisors assessed both the basic skills and the personal traits of applicants. An interviewer must have basic reading and computational skills and be able to follow instructions. Desirable personality traits included receptivity to other's ideas, open-mindedness, and motivation. Additionally, a respondent's willingness to grant an interview often depends on the respondent's initial perception of the interviewer. This perception is influenced by the interviewer's appearance. Thus, candidates who were extreme in appearance and who would not appear "neutral" to any portion of the target population were not selected.

Of the 456 interviewers recruited, two did not attend training and two were released at training.

The characteristics of the 452 interviewers are shown in Table 5-1. More than one-half (260, or 57.5 percent) had worked previously for Westat, and a similar number (271, or 60.0 percent) had worked as interviewers for other field organizations. In addition, 119 of them (26.3 percent) had worked as interviewing supervisors on other, less demanding projects.

The interviewers were primarily middle aged and "young-older," with most (320, or 70.8 percent) between the ages of 31 and 59, only a small number (26, or 5.8 percent) under 30, and slightly less than one-quarter of them (106, or 23.4 percent) age 60 and older. Like most interviewing staffs, the majority of survey interviewers (341, or 75.4 percent) were female.

Table 5-1. Survey research experience, age, sex, schooling, and language skills of the National Adult Literacy Survey interviewers

	Number	Percentage
Westat experience		
Yes	260	57.5
No	192	42.5
Total	452	100.0
Non-Westat experience as interviewer		
Yes	271	60.0
No	181	40.0
Total	452	100.0
Experience as interview supervisor		
Yes	119	26.3
No	333	73.7
Total	452	100.0
Age		
30 or younger	26	5.8
31 to 45 years old	158	35.0
46 to 59 years old	162	35.8
60 or more	106	23.4
Total	452	100.0
Sex		
Female	341	75.4
Male	111	24.6
Total	452	100.0
Schooling		
High school graduate	72	15.9
Vocational degree	23	5.1
Some college	139	30.8
College graduate	92	20.3
Graduate work	25	5.5
Professional degree	47	10.4
Unknown	54	12.0
Total	452	100.0
Spanish-speaking bilingual		
Yes	37	8.2
No	415	91.8
Total	452	100.0

The overwhelming majority of interviewers had some education beyond high school (326, or 81.9 percent), while 164 (41.2 percent) were college graduates. Of the remaining interviewers, "some college" was the mode (139, or 30.8 percent), while level of education was unavailable for 54 (12.0 percent) interviewers.

When Spanish-speaking-only respondents were encountered, the screener and the background questionnaire were administered in Spanish. The field organization included 37 (8.2 percent) interviewers who were bilingual in Spanish and English.

Interviewer attrition on this survey was not excessive for a study of this size or with a field period of this length. Different approaches were used to deal with attrition problems depending on when and where they occurred. In some cases, new interviewers were hired and trained. In other cases, other interviewers working in the PSU were able to complete the remaining work, or interviewers from other areas traveled to the PSUs where the attrition had occurred.

5.4.2 Training

The following sections describe the training of the field supervisors, interviewers, and data editors. Because the survey required an unusually large field staff, a major challenge in planning the training program was to identify the optimal size and scheduling of training sessions. If several relatively small training sessions were held sequentially, the training schedule would have to be extended well into the field period, assuming that it was desirable for interviewers to start work soon after completing their training. This approach would have kept the project management staff out of the home office for an extended period and would almost certainly have jeopardized the productivity and quality of the first weeks of the field effort. Alternatively, scheduling a few very large training sessions presented serious problems of management and limited interviewer practice, thereby jeopardizing the quality of the training program.

The training plan adopted involved eight separate sessions, four at a time in each of two sites. Hence, training was accomplished over a period of 11 days, with an average of 55 interviewers per group. Each group was led by one of Westat's home-office field managers and was supported by the field supervisory staff. The Westat project director and field director and ETS staff monitored all eight sessions.

Another challenge of the training plan was to prepare field staff to act as both survey interviewers and literacy exercise administrators. As is typical, interviewers were trained to take a very active role in conducting the screener and the background interviews and to be prepared to answer any questions the respondent might raise. In the role of exercise administrator, on the other hand, the interviewers had to remain very much in the background, observing and facilitating but intervening only at certain well-defined points and refraining from offering help even if it was requested.

5.4.2.1 Supervisor training

The training material for supervisors was based on the information contained in two manuals—the supervisor's manual and the Automated Survey Control System (ASCS) manual. The supervisor's manual documented and provided instruction on the following supervisory functions:

- Responsibilities before the field work begins, including setting up the office, making initial interviewer assignments, and preparing for training;
- Making additional interviewer assignments and performing the record-keeping tasks associated with assignment preparation;
- Reassigning cases and setting goals for meeting response rates;
- Processing completed cases received from the field;
- Performing field observations and validations;
- Monitoring field progress and costs using reports from the ASCS; and
- Handling problem situations (e.g., interviewers who fail to report on time; have low production rates, high costs, or high refusal rates; or have submitted fraudulent work).

The ASCS is a proprietary Westat software package designed to enable project management and field supervisory staff to maintain close control over field production and costs. Each supervisor was provided with a computer terminal loaded with a file of the records of all households selected for screening in the region. The supervisor updated the household records regularly and transmitted updated cost and care status information to Westat weekly. The ASCS manual included step-by-step instructions for setting up the hardware and connecting the printer and modem line, using the software, and communicating with Westat's home office via electronic mail.

Supervisor training was conducted by Westat project staff in January 1992. The first 2 days of training were a simulation of the interviewer training program (described in detail in section 5.4.2.2), but the smaller size and greater experience of the group made it possible to accelerate the pace. This simulation of the interviewer training program not only prepared the supervisors for their subsequent responsibilities but also provided a dress rehearsal for staff and an opportunity to evaluate and refine materials in preparation for interviewer training. After completing the interviewer training program, supervisors were trained to lead small groups of interviewers through scripted, interactive reviews of the data collection instruments. They were also given training on the supervisor's responsibilities before and during the data collection phase, as described in the bulleted list above.

After completing their training, supervisors returned home to prepare for interviewer training and the assumption of supervisory responsibility immediately after interviewer training.

5.4.2.2 Interviewer training

Basic interviewer training for the 452 interviewers and the 24 data editors consisted of a 3-day, in-person training program, preceded by home study. Novice interviewers received an additional 5 hours of in-person training on general interviewing techniques. Interviewers selected to work on the survey of the prison population were trained in procedures unique to that assignment (described in detail in Chapter 6).

One-half of the interviewers were trained in San Francisco and the remainder in Dallas in early February 1992. Four simultaneous training sessions were held at each site, with approximately 55 trainees per session. One of the four field managers served as the lead trainer for each training group of 55 trainees and was assisted by three field supervisors. Supervisors were assigned to the training session attended by the interviewers from their region. Holding several simultaneous sessions in a site allowed the Westat project director and field director and ETS staff to observe all four sessions at each site.

In April 1992, 35 interviewers were recruited and trained at a small session held in Newark, New Jersey. The majority of the trainees at this session were from northern New Jersey and New York City, where the greatest attrition had occurred. The training program, which was led by the field director and one field manager, was identical to the program used at the initial interviewer training sessions.

Interviewer Training Materials. The training materials were very carefully scripted to cover every concept that the interviewers needed to know, and the scripts were organized into trainer guides. The elaborate preparation of training materials accomplished two purposes. First, it achieved standardization, which is particularly important when a large staff of interviewers is being trained in separate sessions. Second, it allowed all trainers to study the trainer guide, to rehearse their roles, and to be completely prepared for training. This is particularly important in training efforts that require the use of a large training staff. The scripted materials eliminate the necessity for the trainer to improvise. This preparation allowed the National Adult Literacy Survey training sessions to move smoothly and on schedule, which gave the interviewers the confidence that they were being trained by knowledgeable people.

Interviewer Training Techniques. Westat's general approach to interviewer training is centered around five basic training techniques that have been extensively used and refined by Westat survey operations staff over the past 20 years. The following paragraphs briefly describe the five techniques and how they were used for training on the National Adult Literacy Survey.

1. Home study. About 2 weeks before training, the interviewer is asked to read a programmed learning text on general interviewing procedures and to complete practice exercises. The interviewer brings the manual and completed exercises to training, where the field supervisor reviews them and provides feedback.

2. Interactive lecture. This technique is used to provide the first introduction to the questionnaire. The lead trainer uses a scripted lecture to present the basic concepts of the questionnaire to the entire group of trainees. Trainees take turns playing the role of interviewer and asking the questions, while the lead trainer provides responses from the script, making some general points about the questionnaire as the lecture proceeds. Each trainee records the responses in a copy of the questionnaire. A transparency of each page in the questionnaire is projected on a screen in front of the group, and a member of the training staff demonstrates the correct recording of responses. Trainees are instructed to check their own recording against the recording on the screen. Interactive lectures were used for the initial presentations of the survey screener, the background questionnaire, and the exercise booklet. After the procedures for administering the literacy exercise had been introduced in an interactive lecture, a demonstration of the exercise was presented by the lead trainer, with a field supervisor acting as the respondent. The intention of the demonstration was to give trainees a better understanding of their role as exercise facilitator.

3. Mock interviews. After the trainees receive a general introduction to the questionnaire in the interactive lecture, more complicated examples and instructions are presented by the field supervisors to small groups, or "learning communities," of about 18 trainees. Trainees take the role of interviewer, while the field supervisor acts the role of respondent, following a scripted questionnaire. The leader's script includes instructions to interrupt the script at appropriate times to review certain sections of the interviewer's manual, point out some of the less obvious features of the questionnaire, or explain certain terms. The scripts used in the learning communities present increasingly complicated scenarios, so that trainees become familiar with the various types of cases they will encounter.

4. Practice exercises. Written exercises are designed to reinforce and test comprehension of certain concepts. They are particularly well-suited for evaluating the trainee's comprehension of some of the more complicated questionnaire issues. Completed exercises are reviewed by the leader, who gives further instruction to trainees who appear to be having difficulties with the concepts.

5. Dyad role playing. The purpose of role playing is to provide additional practice and to allow the trainee to get a feeling for the overall flow of the interview. Trainees are arranged in pairs. One member of each pair is given a scripted copy of the questionnaire and plays the role of respondent, while the other trainee conducts the interview. With the next script, the members of the pair reverse roles. For the National Adult Literacy Survey, four role-playing scripts were used. The scripts began with the screener and ended with the administration of the literacy exercises.

In-person training program for interviewers. Most of the 3-day interviewer training was devoted to teaching procedures for the administration of the data collection instruments—the screener (national and state samples), the background questionnaire, and the exercise booklet. In addition, instruction was provided on gaining respondent cooperation, keeping records of non-response, editing completed work, and completing administrative forms. Table 5-2 presents an overview of the training program.

The study overview was presented by the Westat and ETS project directors. The purpose of the overview was to provide the interviewers with sufficient background information on the study to enable them to speak knowledgeably about it to survey respondents. The overview included a history of literacy studies and information about how the data would be used.

Training interviewers to administer the exercises presented a particular challenge. The role of exercise administrator is different in important ways from that of survey interviewer, and the interviewers had to learn to switch from one to the other. During the administration of a survey questionnaire, a very dynamic interaction takes place between the interviewer and the respondent. Although the interviewer must remain neutral and avoid leading the respondent, the interviewer provides reassurance and encouragement and shapes the respondent's behavior by giving continual feedback. The administration of the assessment exercises, on the other hand, requires the interviewer to take a much more passive role, observing the respondent's performance without intervening and studiously avoiding any temptation to provide assistance, even when help is requested.

The issues involved in making the switch from interviewer to exercise administrator were discussed in an interactive lecture conducted by a lead trainer. In addition to instructing interviewers on the mechanics of administering the exercise, time was spent discussing problems that might occur in the administration of the exercise, such as respondents with limited English-speaking abilities, respondents with physical or mental conditions that might affect their performance on the exercise, and the special needs of the elderly population. This session ended with a demonstration of exercise administration by a member of the training staff. Then trainees practiced the administration of the exercise in their learning communities and reviewed record-keeping procedures specific to the use of the exercise booklets.

Table 5-2. Overview of the National Adult Literacy Survey interview training session

Day	**Topic**	**Presentation mode**
1	Overview of the study	Plenary session*
	Overview of interviewer's job	Interactive lecture+
	Survey materials	Interactive lecture
	Procedures for gaining respondent cooperation	Learning community**
	Screener	Interactive lecture
	Screener	Learning community
2	Background questionnaire	Interactive lecture
	Background questionnaire	Learning community
	Exercise booklet	Interactive lecture demonstration
	Exercise booklet	Learning community
3	Administrative procedures	Interactive lecture
	Screener, background questionnaire, and exercise booklet	Dyad role playing
	Reporting	Learning community
	Closing lecture on neutrality	Interactive lecture

* All trainees at site in attendance.

+ Fifty-five trainees at each site in attendance.

** Eighteen trainees in each learning community in attendance.

At each training site, the Westat home office staff and field supervisors attending training met in the evening to discuss any problems that had arisen. Minor modifications to the training program or schedule were discussed as a group. Any necessary changes to the materials were made and distributed to the training staff.

Trainees with potential performance problems were identified and remedial measures were discussed at these nightly meetings. Such trainees were closely observed and were paired during role plays with a staff member who could assist them during the role play. One-on-one conversations were held with these trainees regarding their progress. At the end of training, two trainees who were unable to master the procedures and techniques required for the job were released from the study.

The trainee group included 37 Spanish-speaking interviewers. At each session, they were assembled into one learning community, which was led by one of two Spanish-speaking field supervisors. This gave them an opportunity to work with the Spanish translations of the screener and the background questionnaire and allowed the supervisors to assess the Spanish-speaking abilities of the bilingual interviewers.

Two special training sessions were conducted: a half-day session on general interviewing techniques, held the day before the main program began, for 189 inexperienced interviewers; and a 1-day session on special procedures for the prison sample, held the day after the main program, for 51 interviewers assigned to work on that sample.

5.4.2.3 Editor training

Following interviewer training, data editors were trained at a 1-day session also attended by the regional supervisors. The Westat field manager in charge of editing procedures was the lead trainer and was assisted by the ETS staff member who oversaw the editing and coding operation at ETS. Editors were trained on general editing procedures, detailed specifications for editing each questionnaire, data coding of the screener, and data retrieval procedures. Training included practice sessions in which editors and supervisors edited and coded scripted cases.

5.5 FIELD OPERATIONS

The National Adult Literacy Survey field period began in February 1992, immediately following the completion of the first interviewer training session, and lasted for 28 weeks, until the end of August 1992. All components of the survey sample were worked simultaneously, including the national sample (incentive cases and non-incentive cases[2]), the state sample, and the prison sample.

The following sections describe the survey field operations, including the general approach, the schedule and production, and the reporting systems used to manage the effort.

---

[2] See section 2.3 for a discussion of the non-incentive sample.

5.5.1 General Approach to the Field Effort

For the National Adult Literacy Survey field effort, Westat used an approach that has been effective for many previous surveys involving large, complex, in-person data collection operations. Under this approach, the field effort occurs in three overlapping stages:

- *Initial phase:* Each area segment[3] is assigned by the regional supervisor to an interviewer, who follows certain rules in making a prescribed number of calls to every sampled dwelling unit in the segment;
- *Reassignment phase:* Cases that did not result in completed interviews during the initial phase are reviewed by the regional supervisor, and a subset are selected for reassignment (to another interviewer in the same PSU or an interviewer from a nearby PSU); and
- *Special non-response conversion phase:* The home office assembles a special traveling team of the most experienced or productive interviewers to perform a non-response conversion effort, under the supervision of a subset of the field supervisors.

The assignments in the initial phase are controlled by the regional field supervisor. In the National Adult Literacy Survey, the supervisors had two or three local interviewers available in most PSUs. Each area segment was assigned to one of the interviewers based on the racial/ethnic composition of the area and the proximity of the segment to the interviewer's home.

During the initial phase, the interviewers were instructed to make up to four in-person calls to the household to complete a screener and up to four additional in-person contacts -- after completing the screener—to administer the background questionnaire and the exercise booklet, which had to be completed during the same visit. To maximize the chances of finding respondents at home, most contacts were made during prime interviewing hours (3 p.m. to 9 p.m. on weekdays and 10 a.m. to 9 p.m. on Saturdays and Sundays). Contacts at each dwelling unit were to be staggered on different days of the week and at different times of the day. All calls to complete the screener had to be made in person. If the screener was completed and the background questionnaire and exercise booklet could not be completed on the same visit, the interviewer was permitted to use the telephone to set an appointment to administer these questionnaires in person. The initial phase was considered complete when the interviewer reported a definitive outcome for the case or when the full complement of calls had been made.

Interviewers mailed completed cases to the supervisor twice a week. The entire segment folder and all its associated materials were mailed when the initial effort had been completed for all cases in the segment. At the time the assignment was made, depending on the size of the segment, the interviewer was given 2 to 3 weeks to complete the initial effort for all cases in the segment. More productive interviewers

---

[3] Census blocks or groups of blocks within primary sampling units (PSUs).

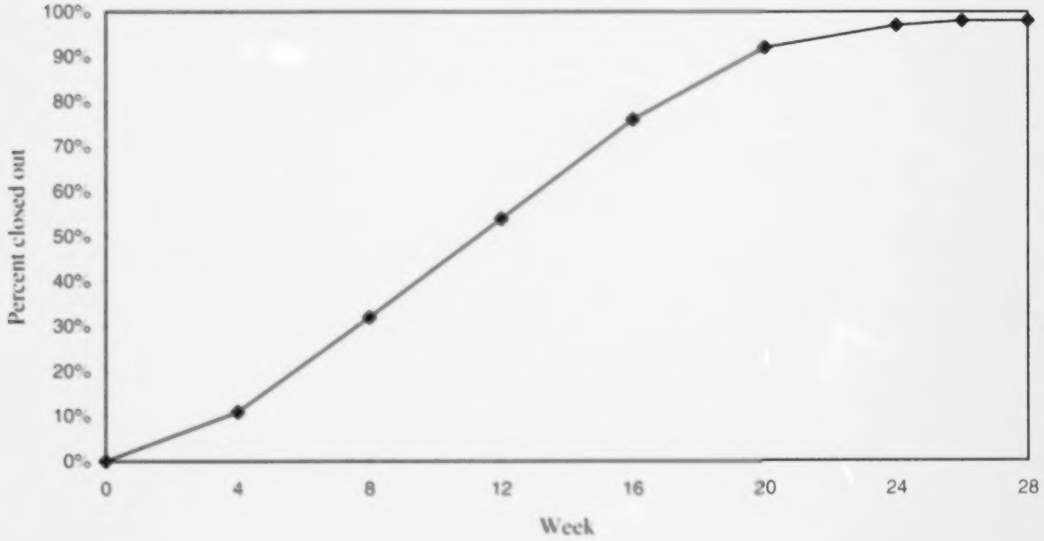
were able to handle up to five segments simultaneously during the initial phase. Less productive interviewers were given only two or three segments at a time.

5.5.2 Schedule and Production

The original plan for the National Adult Literacy Survey field effort envisioned a 24-week field period, in which 20 weeks would be used to complete the initial complement of calls to all assigned households and 4 weeks would be reserved at the end for intensive non-response conversion by the traveling team of interviewers. The second phase was planned to overlap with the first, beginning at about week 13.

In fact, to allow more time to increase response rates, a decision was made to extend the field period to 28 weeks. Figure 5-1 shows the cumulative percentage of survey cases that were closed-out, by month of the field period, for all household sample cases:

Figure 5-1. Percentage of closed-out cases by week of field period



The pace of the field effort was influenced by several factors that were unique to the National Adult Literacy Survey. First, the design called for respondents to complete the background questionnaire and the exercise booklet in the same visit; thus, it was necessary for a respondent to have a period of more than one hour during which the respondent was reasonably unlikely to be interrupted. This requirement reduced the likelihood that respondents would be available on the interviewer's first visit and necessitated additional callbacks to ensure completion of the case. In addition, because of the very large number of cases, across-the-board decisions to improve productivity were cumbersome and took home office staff considerable time to implement. Also, the level of editing and data handling performed in the supervisor's office was unusually high compared to most surveys, where the bulk of the editing task is conducted at the home. This activity may have diverted the regional supervisors' attention from issues of production to a greater extent than was envisioned at the outset.

5.5.3 Reporting Systems

The smooth progress of field work depended on the ongoing monitoring of the interviewers' work and regular communication among the field staff, the Westat home office, Educational Testing Service, and the National Center for Education Statistics. The following sections describe the major mechanisms and procedures used for reporting during the National Adult Literacy Survey field period.

5.5.3.1 Automated Survey Control System (ASCS)

Westat's proprietary field management system, the ASCS, was used to manage and monitor the progress of the field work. This system was designed to utilize a relational database called the survey control file, which contains case records, interviewer records, and cost records. It is the interaction of these records that enables the system to provide information at the level of an individual case or interviewer or to summarize results by aggregating geographically, across time periods, or along other dimensions. The system is thus able to provide critical management information to both supervisors and home office staff.

Each survey supervisor was linked to Westat's home office VAX computer and was able to transmit data to the home office. As an area segment was assigned to an interviewer, the supervisor entered information pertinent to the assignment and the interviewer into a laptop computer. As closed-out cases were received from interviewers in the supervisor's office, result codes were entered into the system. The respondent's race code was also entered for each case so that the sample yield for minority populations could be monitored. The number of hours each interviewer worked and the interviewer's expenses (mileage, tolls, parking, etc.) were entered into the system on a weekly basis. As closed-out cases were edited and validated, editing scores and validation results were entered as well.

Once each week, the supervisor transmitted a copy of the regional database via modem to the home office VAX. After the link, the supervisor could run reports from the system showing the following:

- Productivity and costs by interviewer, weekly and cumulatively;
- Response rates, refusal rates, and percentage of cases closed out at the interviewer, PSU, and regional levels for each survey instrument (screener, background questionnaire, and exercise booklet);
- Editing results by interviewer; and
- Validation results by interviewer.

At the home office, reports were run weekly to monitor sample yield, response rates and completion rates for each type of instrument, number of cases assigned, cost and level of effort, editing results, and validation results. These reports were run at the regional and PSU level and separately for the national, state, and prison samples. Some reports were run separately for the incentive and the non-incentive samples as well. Key reports were sent weekly to Educational Testing Service and to the National Center for Education Statistics.

Additional reports were run each week to allow the project statisticians to monitor the sample yield for minority populations in the national sample and for age-eligible populations in the state samples. Selected variables from the sample selection file, which carried census race characteristics by segment and by PSU, were merged with production data from the ASCS to allow comparison of projected and actual results. These reports were run separately for all cases, incentive cases only, and non-incentive cases.

5.5.3.2 Interviewer reports to the supervisor

Survey interviewers were required to mail closed-out cases twice a week and time and expense reports once a week. They were also required to contact their supervisors by telephone at a regularly scheduled time once a week to discuss all aspects of their work (response rates, production and cost performance, and quality control results). Each outstanding case in the interviewer's assignment was reviewed and discussed. If the interviewer reported that the case had been closed out and mailed, the supervisor noted the reported date of mailing. All assignments were released with an expected completion date, and the interviewers were asked to restate on a weekly basis whether they expected to meet the deadline. The supervisor and interviewer discussed any problems reflected in the ASCS reports (e.g., low response rates, high number of hours per complete case, high error rate found in editing).

An important part of the supervisor's job was determining the optimal flow of work to each interviewer. Based on the weekly conference, the supervisor decided whether the interviewer was ready for an additional assignment. Supervisors tried to maintain a balance between somewhat competing goals—keeping interviewers supplied with enough work to stay productive and not allowing cases to languish by giving an interviewer more work than he or she could close out in 2 or 3 weeks.

5.5.3.3 Supervisor reports to the home office

At least once a week, each supervisor had a telephone conference with the supervisor's home office field manager to discuss progress in the region. Discussion centered on the contents of the week's ASCS reports as well as on current progress as reported to the supervisor during the interviewers' weekly calls. The weekly conferences between field managers and supervisors were used to discuss problems in the region, the prospects and plans for completing the remaining work, and what help, if any, the supervisor needed to complete all work in the region by the end of the field period. The results of quality control procedures were also discussed. If the quality control reports indicated problems with the quality of an interviewer's work, appropriate steps to correct the problem were discussed.

Once a week, a home office staff meeting was held with the project director, the field director, and the field managers to discuss information obtained from the field managers' telephone conferences with the supervisors. The results for each region were reviewed, and any study-wide problems that had developed, such as common editing problems, the interviewing schedule for the prison sample, ASCS software or hardware problems, and distribution of supplies and materials, were reviewed. Strategies for solving problems were discussed and subsequently implemented by the field managers and other staff.

5.5.3.4 Home office staff reports to ETS and to NCES

Each week, a subset of the home office ASCS reports that summarized the progress of the field effort were sent to the ETS Project Director and to the NCES Project Officer. Weekly telephone conferences were held between staff at Westat and ETS to discuss field progress and to review any problems uncovered during the review of the incoming work at ETS. Any important changes in the field work strategy were discussed before implementation. The open lines of communication between Westat and ETS were key to the success of the effort and contributed significantly to the quality and timeliness of the results.

Key staff from ETS and the NCES Project Officer reviewed the field effort on a continuing basis. Major decisions were based on information available from the reporting systems described in this section.

5.6 QUALITY CONTROL OF DATA COLLECTION

5.6.1 Introduction

To ensure that high-quality data were collected for the National Adult Literacy Survey, several quality control measures were implemented, as described below. The procedures were designed to check on the quality and completeness of the data as collected and to provide timely feedback to the supervisors, the home office, and the interviewers.

- A three-part editing program consisting of
    - — an item-by-item edit of completed cases and non-interview reports by the interviewer,
    - — a review of case disposition codes and survey control information by the supervisor, and
    - — an item-by-item edit of each case by the regional office editor;
- Validation of 10 percent of each interviewer's closed-out cases;
- Observation of interviews—in-person and by tape recordings; and
- Observations of supervisors.

Each aspect of the quality control program is described below.

5.6.2 Editing

In the first stage of editing, the field edit, interviewers were required to perform an edit of completed documents. The field edit consisted of an item-by-item review of the case documents for completeness, accuracy, and legibility. Interviewers were encouraged to use the question-by-question specifications in their manual while performing this edit. Corrections on all documents were made using a #2 pencil. On the optically scanned (OPSCAN) documents (background questionnaires and background/exercise non-interview report), any incorrect entries were completely erased and the correct ones were entered. On the screener and the interview guide, erasures were not allowed; interviewers used codes to distinguish between errors that they had made and errors the respondent had made. Any incorrect entries were lined through and the correct entries were recorded. If interviewers discovered that data were missing in any of the documents, they annotated the omission, but they were not permitted to perform data retrieval on their own.

The next stage of quality control occurred in the supervisor's office, where a two-part edit was performed: (1) verification of case disposition codes and other survey control information and (2) a data quality review. First, as each case was received in the regional home office, the supervisor performed a general review, primarily intended to ascertain that all materials related to the case had been returned and that the correct result codes had been entered. During this review, the supervisor entered information about the outcome of the case into the ASCS system.

Second, data editors in each regional office performed a complete data quality edit of all cases. This was a very structured edit in which the editors followed rules in an editor's manual and utilized an edit form for each case to guide them and to document the results. The edit form was divided into sections that listed specific items to be checked for each type of data collection instrument. The completed edit form served as documentation for assigning a final edit score for the case and was also used by the supervisor to give feedback to interviewers about any problems found when their cases were edited.

Editing of completed screener interviews included a review of skip patterns, sampling procedures, and the missed dwelling unit form, as well as an item-by-item review. In addition, editors completed an OPSCAN transcription sheet designed to capture the subset of screener information that would be used in the data analysis. Editing of completed background questionnaires included a review of all entries on the front cover of the instrument, all questions and skip patterns, and the recording conventions of answers. Completed interview guides were reviewed for correct entries on the front cover and completion of the observation section at the end of the document. Editing of non-response cases included a review of the screener non-interview report, or the background questionnaire/exercise booklet non-interview report for completeness and for consistency with the result code assigned to the document.

Nine questions in the background questionnaire were designated key items where correct and complete data were essential for the analysis of the case. If, during the edit, the editor discovered that a key item in a case was missing data or was inconsistent with other questionnaire data, the editor attempted to retrieve the data from the respondent by telephone. On the small number of data retrieval cases where the respondent did not have a telephone, data retrieval was conducted in person, or a decision was made to forego data retrieval.

At the completion of the edit, the editor assigned a final edit result score to the case based on the number of errors discovered, from result score 1 (0 errors) through result score 5 (6 or more errors). The edit scores were entered into the ASCS, and a weekly report was produced showing the edit results for each interviewer.

If an interviewer's edit scores indicated unsatisfactory performance, there were two ways to handle the problem. In cases where it was feasible, supervisors conducted telephone retraining of interviewers whose edit results were unsatisfactory; these interviewers' assignments were also curtailed so that their work could be closely monitored for quality until they had demonstrated improvement. In a few cases, where the problems were considered so great that they could not be remedied by telephone retraining and it was not practical to provide in-person individual retraining, the interviewer was released.

Each regional supervisor reviewed 10 percent of the regional editor's work throughout the field period as a quality control measure. As cases were received and processed at ETS, the ETS staff also provided feedback to Westat on the quality of each editor's work. The evaluation was passed on to the editors in the field.

5.6.3 Validation

A 10 percent subsample of cases was randomly selected for validation by the supervisor. Validation was performed on completed and non-interview cases. For completed cases, the validation interview verified that contact had been made and confirmed the respondent's address at the time of the contact. Then three questions from the background questionnaire were asked again. The respondent was also asked how long the interviewer spent with the respondent on the day of the interview and how much the respondent had been paid for participation in the survey. For non-completes, the validation interview verified contact (if the interviewer's report of the case indicated that contact had occurred), confirmed the respondent's address at the time of contact, and, if possible, tried to schedule an appointment for an interview.

Validation was performed by the supervisors by telephone when possible. On all cases selected for validation for which phone numbers were not available, in-person validation was performed by an interviewer, other than the original interviewer, who worked in the same or a nearby PSU. As soon as validation for a case was completed, the supervisor entered a validation result code for the case into the ASCS.

Because falsification activity that goes undetected for a long time is very costly to correct, it is desirable to perform validation as soon after interviewing as possible. Therefore, if a regional supervisor was unable to keep up with the required validation effort, the validation effort was moved into Westat's home office until the supervisor was able to resume responsibility for performing validation.

Westat home office staff used the ASCS to monitor the progress of the validation effort and to ensure that at least 10 percent of each interviewer's work was being validated. At the end of data collection, a total of 6,068 cases had been validated either by telephone or in person, for an overall validation rate of just over 13 percent. Some of the validation workload in excess of the original 10 percent requirement resulted from situations in which additional verification was needed to dismiss or confirm suspected falsification.

Out of the 452 field interviewers who worked on the National Adult Literacy Survey, seven were discovered to have submitted some fraudulent work. Each of these interviewers was dismissed as soon as falsification was confirmed, and all of their completed work was validated or redone. Three of the seven interviewers were detected and dismissed very early in the field period. A fourth interviewer did not begin to falsify work until near the end of the field period, and the 100 percent verification revealed that the number of her cases falsified was very small. The other three interviewers had each closed out a substantial number of cases before fraudulent work was discovered. When falsification was confirmed, other interviewers conducted in-person validation of all of their cases and discovered that two of the three interviewers had very high falsification rates (70 percent), while the third had a falsification rate of 20 percent. Because these cases were discovered late in the field period, the interviewers completing the falsified cases conducted the re-interviews into the month of September, after the formal shut-down of the data collection effort.

5.6.4 Observation

Two types of observation of the National Adult Literacy Survey interviews occurred—tape recording of interviews for review by supervisors and home office staff and in-person observation, primarily by home office staff.

Tape-recorded interviews. Supervisors relied on review of tape-recorded interviews to "observe" each interviewer. Each interviewer was required to tape record one complete interview, that is, the entire background questionnaire and exercise booklet (to the extent that the respondent was able to do the exercise booklet). The case to be tape recorded was designated in advance by the supervisor, and each interviewer was required to record the case early in the data collection period. After listening to the tape, the supervisor completed a taped observation form and gave feedback to the interviewer on the quality of the interviewing techniques and on any mistakes the interviewer was making.

The supervisors sent the tapes to the home office, where the field managers also listened to many of them as an additional quality control measure. The home office staff also used the taped interviews to learn about respondents' reactions to the survey questions and to get a "feel" for how field procedures were working.

In-person observations. Interviewer observations were also performed by the home office field managers, other members of the Westat and ETS staffs, and by supervisors whose field offices were in sampled PSUs. Interviewer observations were performed for two main purposes. One purpose was to provide home office staff with an opportunity to observe respondents' reactions to the survey and also to observe how well field procedures worked. Supervisors identified their strongest interviewers for this type of observation. The second purpose was to observe interviewers whose performance was of some concern, either because of their evaluation during training or because they were assigned to a particularly difficult area.

Interviewers were typically observed locating sampled dwelling units, making screener contacts, setting appointments, and completing at least one background questionnaire and exercise booklet. During an interview, the observer listened and followed along using a copy of the questionnaire and interview guide, but did not participate in any way. After the interview, when the observer and interviewer had left the respondent's home, the observer used an interviewer observation form to evaluate the quality of the interviewer's work. Interviewers were evaluated on the following points: organization of material, knowledge of the survey, interviewing techniques, and traits such as motivation, perseverance, and tact.

5.6.5 Supervisor Observations

Westat home-office and ETS staffs conducted in-person observations for 9 of the 24 regional supervisors. The supervisors were chosen for observation either because this was their first supervisory assignment for Westat or because the workload in their regions presented unusual problems. The supervisors were evaluated on the following points: office set-up, keeping up-to-date records on production data, data editing, and supervision of interviewers.

5.7 RESPONSE RATES

Response rates were calculated separately for the national and state samples for the screener, the background questionnaire, and the exercise booklet. This section describes the response rates at each of these levels. The results reported for the national survey only include respondents who were paid a $20 incentive. (A detailed discussion of the national survey response rate comparison for incentive and non-incentive cases can be found in Chapter 10.)

The screener response rate was calculated as follows:

$$\text{Screener response rate} = \frac{\text{Total completed screener interviews}}{[(\text{Total sample households}) + (\text{missed DUs}) + (\text{nonexistent or vacant DUs})]} \qquad (1)$$

where "total completed" means that all items had to have a valid response.

A screener was considered complete if enough information had been obtained to select respondents for the background questionnaire and exercise booklet. Therefore, the numerator includes completed screeners for households in which respondents were selected and for households that contained no eligible respondents[4]. Study procedures required that screening information be obtained from a household member who was 16 years of age or older. Toward the end of the field period, interviewers were given permission to complete a very small number of screeners with a neighbor as long as a confirmation of the information could be obtained from a second neighbor, landlord, or relative of a household.

The background questionnaire (BQ) response rate was calculated as follows:

$$\text{BQ response rate} = \frac{\text{Total completed BQ interviews}}{\text{Total number of respondents eligible for the BQ}} \qquad (2)$$

where "completed" is defined as all items had to have a valid response.

To be a "completed" exercise booklet, the following criteria had to be met:

1) Every item in the exercise booklet has an entry or mark. OR
2) Every item in the core section of the exercise booklet has an entry or mark AND at least 5 items have an entry or mark in each of the three blocks of exercises. OR
3) Every item in the core section of the exercise booklet has an entry or mark AND every item in at least one of the blocks has an entry or mark.

A partially completed exercise was one that had an entry or mark for at least one item in the exercise booklet, but did not meet the criteria of a complete (as described above).

Persons eligible to complete the background questionnaire were those selected at the time of screening. One person was selected in households with fewer than four persons eligible for selection, while two persons were selected from households with four or more eligible members.

The exercise booklet (EX) response rate was calculated as follows:

$$\text{EX response rate} = \frac{\text{Total completed exercise booklets}}{[(\text{Total number of respondents who completed the BQ}) - (\text{exercise booklet ineligibles})]} \qquad (3)$$

where "completed" is defined as all items had to have a valid response.

"Exercise booklet ineligibles" were persons who partially completed or could not begin the exercise booklet because they were unable to read English, had a physical or mental disability (such as a learning disability, mental or emotional condition, mental retardation, hearing or visual impairment, speech or language impairment, or a physical disability) that prevented them from completing an exercise, or were unable to read or write.

---

[4] Households in which a screener was completed but no eligible respondents were found included national sample households that were non-minority in high-minority segments and state sample households with no household member age 16-64. (See chapter 2 for a more detailed discussion.)

Because the differences between weighted and unweighted response rates were small, this chapter presents only unweighted response rates. Weighted response rates are provided in Chapter 3.

5.7.1 Reasons for Non-response

The reasons for non-response for the national and state samples are shown in Tables 5-3 and 5-4. Table 5-3 shows response rates for the screener and background questionnaire. No dramatic differences appear between the national and state samples. In the two samples combined, a screener was completed at 34,193 of the 38,378 occupied housing units in the sample, for an overall screening response rate of 89.1 percent. The largest category of screener non-response was "refusal or breakoff," which occurred in 2,381, or 6.2 percent, of the occupied households; the second largest category was "not at home after maximum calls," which occurred in 1,100, or 2.9 percent, of the occupied households; and there were 704 cases of non-response (1.8 percent) classified as language problems or "other," (including illness, disability, or unavailability during the field period).

Table 5-3. The National Adult Literacy Survey national (incentive-only cases) and state sample responserates: Screener and background questionnaire

	National		State		Combined	
	Number	Percent	Number	Percent	Number	Percent
Screener						
All sampled dwelling units	22,989		20,794		43,783	
Occupied dwelling units	20,034		18,344		38,378	
Incompletes	2,246	11.2%	1,939	10.6%	4,185	10.9%
Refusal/breakoff	1,242	6.2%	1,139	6.2%	2,381	6.2%
Language problem	68	0.3%	43	0.2%	111	0.3%
Not home after max calls	575	2.9%	525	2.9%	1,100	2.9%
Other non-response*	361	1.8%	232	1.3%	593	1.5%
Completed screener interview	17,788	**88.8%**	16,405	**89.4%**	34,193	**89.1%**
Background questionnaire						
Eligible respondents	16,590		14,216		30,806	
Incompletes	3,003	18.1%	2,864	20.1%	5,867	19.0%
Partial complete/breakoff	31	0.2%	12	0.1%	43	0.1%
Refusal	1,801	10.9%	1,866	13.1%	3,667	11.9%
Language problem	100	0.6%	114	0.8%	214	0.7%
Physical/mental disability	320	1.9%	119	0.8%	439	1.4%
Not home after max calls	462	2.8%	480	3.4%	942	3.1%
Other non-response	289	1.7%	273	1.9%	562	1.8%
Completed interviews	13,587	**81.9%**	11,352	**79.9%**	24,939	**81.0%**

* Illness or disability, unavailable during field period, etc.

The screening effort identified a total of 30,806 eligible respondents, of whom 24,939, or 81.0 percent, completed the background questionnaire. Refusals and breakoffs, the largest category of non-response to the background questionnaire, occurred in 3,710 cases (3,667 refusals and 43 breakoffs), accounting for 12 percent of all eligible cases. The next largest category was "not at home after maximum calls," which accounted for 942 cases, or 3.1 percent. There were 439 persons who were considered non-respondents as a result of a physical or mental disability, accounting for 1.4 percent. All other non-response, including language problems and persons who were unavailable during the field period, accounted for 776 cases, or 2.5 percent.

Table 5-4. The National Adult Literacy Survey national (incentive-only cases) and state sample response rates: Exercise booklet

	National		State		Combined	
	Number	Percent	Number	Percent	Number	Percent
Exercise booklet						
Eligible respondents	13,591		11,353		24,944	
Incompletes	1,961	14.4%	876	7.7%	2,837	11.4%
Partially complete	1,029	7.6%	488	4.3%	1,517	6.1%
Language barrier	223	1.7%	117	1.0%	340	1.4%
Physical/mental disability	232	1.7%	66	0.6%	298	1.2%
Reading/writing barrier	316	2.3%	141	1.2%	457	1.8%
Breakoff, unwilling to continue	79	0.6%	63	0.6%	142	0.6%
Other non-response	179	1.3%	101	0.9%	280	1.1%
Totally incomplete	932	6.8%	388	3.4%	1,320	5.3%
Language barrier	343	2.5%	107	0.9%	450	1.8%
Physical/mental disability	191	1.4%	34	0.3%	225	0.9%
Reading/writing barrier	79	0.6%	31	0.3%	110	0.4%
Refusal	290	2.1%	199	1.8%	489	2.0%
Other non-response	29	0.2%	17	0.1%	46	0.2%
Completed exercises	11,630	**85.6%**	10,477	**92.3%**	22,107	**88.6%**
Response rate[5]		**95.3%**		**96.5%**		**95.8%**

Table 5-4 shows the reasons for non-response to the exercise booklet. A total of 24,944 persons were classified as eligible for the exercises. Of these, 22,107, or 88.6 percent, completed the booklet and an additional 1,517, or 6.1 percent, partially completed it. The main reasons for partial completion were reading or writing barriers (457), language barriers (340), and physical or mental disability (298). Of the 1,320 persons who did not attempt the exercise booklet, 489, or 2 percent of all eligible persons, were classified as refusals; 450, or 1.8 percent, were classified as having a language barrier; while the remaining 381, or 1.5 percent, were classified as having a physical or mental disability (225), as having a reading/writing barrier (110), or as non-respondents for some "other reason" (46).

5.7.2 Characteristics of Non-respondents

Tables 5-5 and 5-6 present a breakdown of response rates for the background questionnaire and the exercise booklet by age, sex, and race/ethnicity categories. Response rates are shown separately for the national sample, the state sample, and the two samples combined.

---

[5] The denominator used in calculating the response rate excludes those who had language barriers, physical/mental

Table 5-5. The National Adult Literacy Survey background questionnaire response rates by respondent characteristics

	National	State	Combined
All respondents	81.9	79.9	81.0
Age	86.5	83.4	85.0
16-24	84.4	81.2	82.8
25-44	80.4	77.3	78.7
45-64	77.4	—	77.4
65+	137	74	211
Missing data (N)			
Sex			
Male	79.6	76.1	77.9
Female	83.8	83.2	83.5
Missing data (N)	15	14	29
Race/ethnicity			
Hispanic	81.6	81.9	81.7
Black, non-Hispanic	85.4	82.8	84.6
White and other	80.9	79.5	80.2
Missing data (N)	24	39	63

The response rate results for both the background questionnaire and the exercise booklet exhibited a distinct age differential, that is, younger respondents were considerably more likely to respond to both. Since this is not typically the case in surveys, the most obvious explanation would seem to be that the idea of completing a set of exercises testing reading and writing skills was more threatening to respondents who were further away from school age. Another possible explanation is that the response rate for younger respondents was more favorably affected by the monetary incentives.

A strong gender effect can also be seen in the response rates for the background questionnaire. Women had a markedly higher response rate for this questionnaire, which is consistent with experience on many other survey efforts. However, having completed the background questionnaire, men and women were about equally willing to complete the exercise booklet.

Race/ethnicity seemed to have a moderate impact on response rates to the background questionnaire, with black, non-Hispanic respondents having the highest response rate (84.6 percent), Hispanic respondents having the next highest rate (81.7 percent), and White and all other respondents the lowest rate (80.2 percent). However, the exercise booklet response rates exhibited a slight reversal of this trend, with the White and other category having the highest rate (96.3 percent), Hispanic the next highest rate (95.0 percent), and black, non-Hispanic the lowest response rate (94.3 percent).

---

disabilities, or reading/writing barriers of both the partially completed and totally incompleted categories.

Table 5-6. The National Adult Literacy Survey exercise booklet response rates by respondent characteristics

	National	State	Combined
All respondents	95.3	96.5	95.8
Age			
16-24	98.4	98.0	98.2
25-44	96.3	97.0	96.7
45-64	94.5	94.7	94.6
65+	88.9	—	88.9
Sex			
Male	95.1	96.4	95.7
Female	95.4	96.6	96.0
Race/ethnicity			
Hispanic	95.1	94.7	95.0
Black, non-Hispanic	94.1	94.7	94.3
White and other	95.7	96.9	96.3
Education level			
Some or no high school	92.5	96.2	94.0
High school graduate/GED	95.1	95.7	95.4
Some college or vocational education	96.8	96.5	96.6
College graduate or adv degree	96.4	97.8	97.1

It was not possible to compare the education levels of respondents and non-respondents to the background questionnaire, because the information on respondents' education level was collected on that questionnaire. Not surprisingly, the exercise booklet response rate exhibited a differential with respect to respondent education, although it was not nearly as dramatic a result as the differential that occurred for respondent age.

5.7.3 Discussion

The initial phase of data collection began in early February 1992, immediately after training, and was completed for more than 90 percent of the cases in five months. Initial work on 10 percent of the cases continued for one additional month. During this last month of the initial effort, interviewers from nearby PSUs were sent to close out the initial effort in PSUs where local staff had not been able to do so. Five areas of the country required the efforts of interviewers from other PSUs to complete the initial phase of the work: Dallas/Ft. Worth, Texas; San Francisco/Oakland, California; Newport/Providence, Rhode Island; Queens and the Bronx, New York; and Northern New Jersey. Interviewing teams in these five areas were led by supervisors who traveled to the sites and made interviewing assignments, met with the interviewers on a daily basis, and developed strategies to complete the work quickly, efficiently, and with the highest response rates possible. The camaraderie created by team work and the close, on-site orchestration of the effort by the supervisor made this approach highly successful.

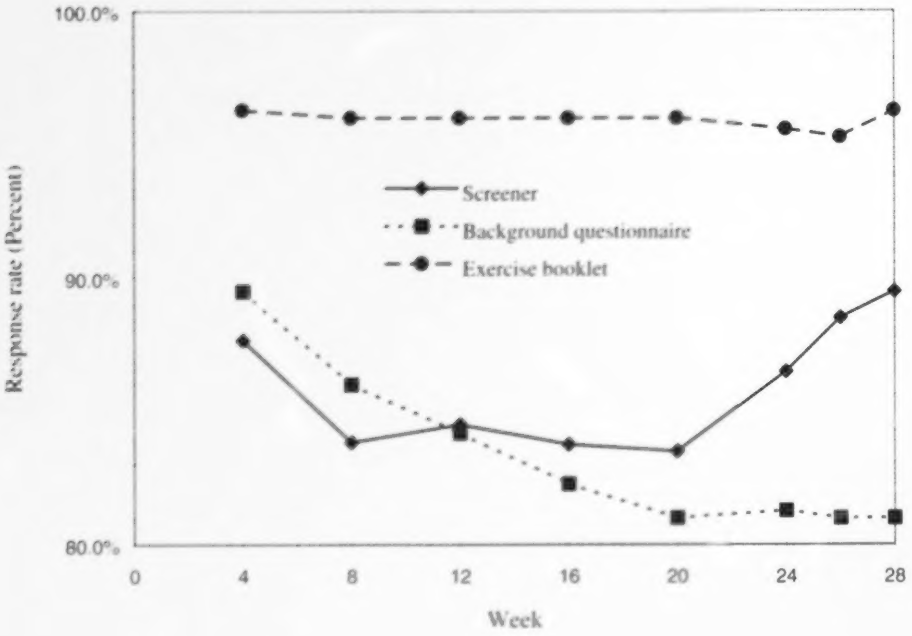
The second phase of the field effort began towards the end of the initial phase and continued through the end of the data collection effort. In this phase, the supervisor reviewed each case that had been closed out as a non-interview during the initial phase. If, in the judgment of the supervisor, there was a reasonable probability that additional effort by the same interviewer or a second interviewer could convert the case, the supervisor reassigned the case accordingly. Judicious reassignment of cases is a very effective technique for increasing response rates, but it is most effective if there is at least one interviewer who has completed the initial assignments promptly, and with good response rates, in the same PSU or a nearby PSU. In such instances, incomplete cases can be reassigned to the more successful interviewers, and response rates are improved. Refusal households selected for reassignment received one of three refusal conversion letters (described in section 5.3.5.2) in advance of the reassignment contact.

The third phase was by far the most expensive and required that the best interviewers in the national field force be identified. For these reasons, this phase was directed from the home office. Preparation for this phase began in the sixth month of data collection. As a first step, supervisors sent all remaining non-response cases to the home office for review by the field managers. Using both the ASCS reports and information gleaned from case-by-case review, the field managers and field director selected sites with the greatest potential for non-response conversion. The field managers and field director were responsible for both case assignments and the formation of traveling teams. In the last month, the field force was consolidated into eight regions.

During phase 3, 16 traveling interviewers covered 22 sites, and a combination of traveling and local staff reworked non-response cases in two sites. This effort raised the screener response rate from 88.1 percent to 89.7 percent (Figure 5-2).

The tasks required of respondents and the time commitment required were significant. The response rates achieved were quite favorable. Further, the interviewers were able to capture detailed reasons for exercise booklet non-response. These data were used to assign an exercise scoring algorithm that took advantage of this information. With additional time and resources, response rates would have been increased. In addition, relying on confirming information from two neighbors to "complete" a household screener is a technique used on many large and visible in-person data collection efforts sponsored by the government. If this procedure had been followed throughout the data collection effort, the screener response rate would have been considerably higher. The payment of a monetary incentive is critical to meeting the response rates achieved on the National Adult Literacy Survey.

Figure 5-2. Response rates by week of field period



Chapter 6

THE PRISON SURVEY

Merle Klein and Martha Berlin, Westat, Inc.

The survey of the prison population was recommended by two committees that advised the National Adult Literacy Survey, the Technical Review Committee and the Literacy Definition Committee. The survey was then funded by the Division of Adult Education and Literacy in the Office of Vocational and Adult Education, U.S. Department of Education and incorporated in the contract between the National Center for Education Statistics and Educational Testing Service to conduct the literacy survey. The prison survey component was developed in consultation with the Bureau of Justice Statistics (BJS) of the U.S. Department of Justice and the Federal Bureau of Prisons (FBOP).

The purpose of the survey was to assess the literacy skills of adult inmates in Federal and state correctional facilities, using a sample size of approximately 1,000 inmates. The sample of 1,000 inmates helped provide better estimates of the literacy levels of the total population and made it possible to report on the proficiencies of this important segment of society. This research was the first in-person literacy assessment involving the prison population.

6.1 Sample Design

A two-stage design was used to select a nationally representative sample of the population incarcerated in Federal and state adult correctional institutions in the 50 states and the District of Columbia.

In the first stage, a sample of 96 facilities was selected (eight Federal prisons and 88 state facilities) from a computerized directory maintained by BJS and FBOP. The prisons were stratified according to the following characteristics: Federal or state prison, region of the country, number of male/female inmates, and level of security.

To select the second stage of the sample—inmates within the sampled facilities—interviewers drew a systematic sample of 9 to 22 inmates from each facility. The lists provided by the prison identified inmates who were assigned a bed and had slept at the facility the night before sample selection. The sample selection procedures were specified in a series of worksheets that the interviewers were required to follow and complete. The interviewer reviewed the worksheets with the field manager by telephone before the prison coordinator was asked to schedule interviews with the prisoners associated with the selected identification numbers. A complete description of the sample design is found in Chapter 2.

To ensure comparability with the National Adult Literacy Survey household component, the literacy tasks administered to inmates were the same as those given to the household population. However, to address issues of particular relevance to the prison population, a revised version of the background questionnaire was developed. The instrument drew questions from the 1991 Survey of Inmates of State Correctional Facilities, sponsored by BJS. These included queries about current offenses, criminal history, and prison work assignments, as well as education and work force experience.

In keeping with the procedures established for the household survey, a monetary or in-kind incentive was to be provided to inmates who participated. However, the rules of almost all facilities precluded such an incentive. Instead, a personalized Certificate of Participation was given to the facility coordinator to be placed in the inmate's file upon completion of the survey.

In advance of the main survey data collection, a small pretest was conducted at a correctional institution in Maryland. The pretest was designed to evaluate the ease of administration of the survey instruments, survey administration time, within-facility procedures, and inmate reaction to the survey. The pretest experience demonstrated that several changes to the background questionnaire would facilitate administration. These were made before the main data collection effort. Administrative procedures were also refined to reflect lessons learned during the pretest.

6.2 Gaining Cooperation

The permission and cooperation of Federal, state, and correctional facility officials was required in advance of interviewing in the prisons. The Advisory Panel to the prison component of the National Adult Literacy Survey was particularly helpful in suggesting approaches to gaining cooperation, in securing letters of endorsement, and in reviewing materials sent to correctional system officials. The panel was composed of representatives from FBOP, BJS, and the Office of Vocational and Adult Education within the U.S. Department of Education. The Westat negotiating team was led by a senior staff member with extensive experience in securing the cooperation of a wide variety of officials and in conducting research studies within correctional facilities. Letters of endorsement were obtained from the Correctional Education Association and the American Correctional Association. All of these organizations and individuals contributed to the success of the negotiation process. Of the 96 facilities originally selected to be included in the study, only 88 were asked to participate because of the favorable responses from the facilities contacted early in the negotiation process.

The following steps were used to gain cooperation at the sampled facilities.

- Westat mailed letters to the state corrections officers in all states with prisons selected for survey participation. A letter was also mailed to officials at FBOP. The letter explained the survey and asked for permission to contact the selected facilities within the officer's jurisdiction. Letters were followed up with telephone calls to answer questions, secure cooperation, and determine prison contact procedures.
- The state or Federal official, in most cases, informed the warden that the facility had been selected and urged participation before Westat called the facility. The warden was asked to approve the study protocol and to designate a prison official to serve as prison coordinator for the survey. The Westat prison coordinator obtained the list of inmates from which a sample was

selected, worked out the interviewing procedures within the facility, and scheduled the interviews.
- The Westat interviewer assigned to conduct interviews at a facility contacted the prison coordinator two days before the scheduled sampling date to reconfirm negotiated arrangements and to resolve any outstanding details.

Of necessity, facility negotiations included (1) procedures for providing interviewer security within the institution and (2) interviewer clearance procedures required by the facility. Prison coordinators were asked to arrange a secure, private room for each interview. If this was not possible, interviews were conducted in partitioned or private areas of larger rooms where the inmate could complete the survey uninterrupted and assured of confidentiality. Depending on the security regulations at each facility, respondents were either brought to the interview session by a guard or were provided with a pass to meet with the interviewer unescorted. To minimize misinformation and deter refusals, facilities were requested to simply "call out" selected inmates without providing an explanation of the survey. The interviewer was responsible for introducing the survey and gaining inmate participation.

To obtain clearance for interviewers to enter the correctional institutions, the interviewer's name, address, Social Security number, date of birth, and driver's license number were submitted to the facility and were typically processed by a recognized clearance agency. Also, survey materials were generally reviewed by prison officials during the negotiation process; in some cases, officials denied permission for interviewers to carry in the tape recorder used to administer some literacy tasks. For those cases, the tasks were administered but not recorded. Most facilities also required that interviewers obtain from inmates a signed informed consent form before the interview. The form included statements on confidentiality and assurances that participation or nonparticipation would not affect release or parole eligibility.

6.3 Interviewer Selection and Training

For the survey of the prison population, 51 interviewers were recruited from among the household survey workforce. Criteria for selection included proximity to sampled facilities, experience in interviewing in correctional facilities, availability for the number of hours required to complete both household and correctional facility assignments, and a willingness to interview in correctional facilities. Interviewing manuals were designed to explain procedures specific to interviewing the prison population. Included were instructions on the use of facility contact and sampling forms, question-by-question specifications for each of the data collection instruments, and reporting information. Specifically, training materials focused on the following:

- The background and purpose of the study, including an overview of facility negotiations;
- Inmate sampling forms and procedures;
- Question specifications for administering the background questionnaire;

- Procedures for working within correctional facilities, obtaining inmate cooperation, and reporting results of the surveys; and
- Administrative procedures.

Interviewer training was conducted in a 1-day, in-person session following the training for household survey interviewers. Particular emphasis was given to inmate sampling and collecting data on criminal history and prison employment. Home study packages were sent to the interviewers several weeks before training. These materials included exercises on sampling procedures similar to those completed during the in-person sessions and scenarios that required the interviewer to record data on criminal offenses.

6.4 Data Collection

On average, five days were required to select the sample of inmates and to administer the required interviews in each facility. Interviewers usually entered the facility on Monday morning to sample inmates and to submit the list of selected identification numbers to the prison coordinator for scheduling interview appointments. Interviewing began on Tuesday and proceeded at the rate of about four interviews per day. Interviewer assignments were guided by the proximity of the interviewer's home to the facility and by interviewer availability.

Because retrieval of missing data was not possible, interviewers were required to perform an initial edit before the inmate left the interviewing room. Key-item questions were designated on the forms so that the interviewer could quickly scan the documents for missing information.

The response rates achieved on this unique effort were quite favorable. Of the 1,340 inmates selected, 1,147 (85.6 percent) completed the background questionnaire. The exercise booklet was completed by 996 inmates and partially completed by another 107. These rates are a significant achievement, especially since interviewers had no control over the availability of selected inmates within the short data collection period at each prison and the ability of inmates to complete the exercises. Although interview appointments were scheduled by the prison coordinator, sessions were sometimes delayed, interrupted, or cancelled due to unscheduled inmate count-downs, facility lockups, or movement of a sampled inmate to solitary confinement, restricted housing, or a unit for the mentally ill. Prisons also changed or cancelled appointments to accommodate inmate obligations, such as attorney meetings or court appearances.

6.5 Quality Control

The measures used to ensure the collection of high-quality data included structured edits and a series of communications between interviewers and supervisors. Each interviewer telephoned the field manager to review the sampling results immediately after completing inmate selection. Any problems were referred to Westat's statistical staff before the interviewer could proceed. At the conclusion of the first day of interviewing, each interviewer was required to telephone the regional supervisor to review the background questionnaires for completeness.

Editing specifications and a screener transcription sheet codebook were specifically prepared for the survey of the prison population. Editors who had worked on the National Adult Literacy Survey household component were provided an additional day of training to familiarize the editors for the inmate component with the requirements and documents specific to the prison interviews.

Chapter 7

PROCESSING THE DATA

Anne Campbell, Diné College (formerly of Educational Testing Service);
Norma Norris, Educational Testing Service

7.1 Receipt Control

When a shipment was received at ETS, staff checked the contents against the enclosed transmittal form, which served as the packing list for each shipment. The transmittal form indicated the name of the supervisor who sent the package, the number of boxes or pieces sent, and the total number of cases enclosed. The documents provided with each case varied, but in order to be considered a valid case, each packet of documents had to include a call record folder and a screener. Other documents each case might include were a background questionnaire, an exercise booklet, and a non-interview report form (used only when either the background questionnaire or exercise booklet was incomplete). The number of respondents selected within a household determined the number of documents found within a call record folder.

At check-in, staff verified the number of call record folders against the number recorded on the transmittal form. Then, each document enclosed in a call record folder was thoroughly checked for consistency. Each call record folder was preassigned a unique 10-digit Westat identification number, signifying a household. Each respondent within a household was identified by adding an eleventh digit to the 10-digit number. All documents enclosed within a call record folder had this 11-digit number and a 6-digit exercise booklet number recorded on them, thereby linking these documents back to the individual who completed them.

7.1.1 Screener

There were two versions of the screener: the national household screener and the state household screener. Both of these screeners were printed in English and Spanish and were readily distinguishable by being printed on different color stock.

For both the national and state screeners, staff verified that:

- The label affixed to the screener containing the Westat number and exercise number(s) matched the label on all of the documents enclosed in the call record folder;
- The time at which the interview began was recorded (if not, 99:99 was recorded to indicate "time unknown");
- The questions on the screener were completed properly; it was necessary to go through the screener question by question, to ensure that the correct respondent was selected for participation in the survey;
- The interviewer used the correct procedures in selecting a respondent when there was a "race message" recorded on the front of the call record folder (this message occurred only in the national sample); and

- The selected respondent was age eligible; for both the national and state samples, the respondents had to be at least 16 years old. The state sample had an age cap of 64, whereas the national sample had no age cap.

If the screener contained an error, the entire call record folder was pulled, and its materials were not processed until the error was resolved. In some cases, such as those where selection procedures were not followed correctly, the case was sent back to Westat, who, in turn, reassigned it to the field.

Each screener contained a data transcription sheet, which was used to summarize the information collected in the screener. Since the screener was too complicated to key enter, certain information was transcribed from the screener onto the scannable data transcription sheet. The data transcription sheet was proofed during check-in. A second staff person performed a 100 percent quality control check, to ensure that critical information was entered and was consistent with the screener. If errors were found in the transcribing of information from the screener onto the data transcription sheet, the errors were corrected before the sheets were sent to the scanning department.

7.1.2 Background Questionnaire

There were three versions of the background questionnaire: national/state, national non-incentive, and national prison. Only the national/state and prison background questionnaires had English and Spanish versions. At check-in, it was necessary to verify that the information transcribed onto the background questionnaire's cover was both written and gridded correctly, including such information as the Westat identification number, the exercise booklet identification number, the background questionnaire elapsed time, and the exercise booklet elapsed time. In cases where either the respondent did not complete the exercise booklet or the elapsed times were omitted by the interviewer, 999 was recorded and gridded.

For each background questionnaire, staff verified that certain questions providing critical information had been answered. These included questions on education level, employment status, parents' levels of education, race, and sex. If a response was missing, the case was returned to Westat for data retrieval.

Information in the background questionnaire and corresponding screener information were verified against each other. Discrepancies were possible, as the person who provided information for the screener may not have been selected as a respondent. Rules were established for resolving discrepancies, and in some instances the cases were sent back to Westat for data retrieval.

7.1.3 Exercise Envelope

Each exercise envelope contained an exercise booklet, an interview guide, and (with the exception of one booklet) a newspaper. There were 26 unique exercise booklets, each containing four blocks of tasks: the core (same for all exercise booklets) and 3 cognitive blocks.

At check-in, staff verified the Westat identification number and paged through the exercise booklet to determine whether it was to be considered complete or partial. A booklet was considered complete if it met one of the following three criteria:

- The entire book, that is, the core and all three of the cognitive blocks, had been completed
- The core and at least one of the three cognitive blocks had been completed
- The core and five tasks in each of the three cognitive blocks had been completed.

Once the screener, background questionnaire, and exercise booklet had been checked in, they were separated and placed into bundles of 26. Each bundle was assigned a unique number and entered into a LOTUS file for tracking purposes.

The screeners were given to a proofer, who did a quality control check of the data transcription sheet against the screener. Once the transcription was verified as correct, the data transcription sheet was separated from the screener and sent to the scanning department for processing. (See Section 7.1.1 for further detail.) The background questionnaires were given to coders who coded the open-ended items, and the exercise booklets were given to readers who scored the open-ended literacy tasks.

7.2 Coding and Scoring

The background questionnaires and the exercise booklets contained open-ended questions or tasks that required coding or scoring. Background questionnaire items that needed to be coded included questions on country of birth, languages spoken, wages, industry and occupation, and date of birth. Nearly all the simulation tasks were open-ended and so required scoring by trained readers.

A group of nine persons coded the background questionnaires, and a group of 11 scored the exercise booklets. The background coders all had at least a bachelor's degree and had a variety of work experiences, including editing, accounting, communications, and historical research. The scorers for the exercise booklets all had at least a bachelor's degree in education, English, mathematics, or journalism. Both groups included men and women of various ages and racial/ethnic backgrounds.

The scoring supervisor monitored both the coding of the questionnaires and the scoring of the exercises. The scoring supervisor reviewed discrepancies between scorers for the exercise booklets. To facilitate the coding of the questionnaires, the supervisor delegated the responsibility of resolving discrepancies between coders to three of the most accurate coders.

7.2.1 Coding Background Questionnaires

Coding guides were drawn up for the questions dealing with country of birth, languages, wages, and date of birth. For example, numerical codes were assigned for countries and languages, and guidelines were established for converting date of birth into numerical codes. Guidelines were also established for converting wage amounts into a weekly wage equivalent. The *Alphabetical Index of Industries and Occupations* was used to code the industry and occupation questions.

Copies of responses to questions requiring coding were made as the first questionnaires were received at ETS. The scoring supervisor used these responses to train the coders. The coding guides were explained to the coders in connection with sample responses, and then the coders coded an additional 10 to 25 responses. The supervisor then went over the codes for the responses and discussed those that presented particular problems. The training for the industry and occupation questions was conducted by an experienced coder from Westat. The trainer introduced the index of codes and discussed sample responses with the coders, who then coded 50 additional responses. The trainer then went over the codes for the responses and explained the rationale used in arriving at the codes.

After the training was completed, the coders commenced the coding of the questionnaires. For several days, the trainer monitored each person's coding of the industry and occupation questions and discussed the coding of responses with each person. In addition, she conducted a follow-up training session. After the training was completed, the trainer identified three coders who were the most accurate in their coding to resolve discrepancies in industry and occupation codes.

In order to monitor the accuracy of the coding, the questions dealing with country of birth, language, wages, and date of birth were checked in 10 percent of the questionnaires by a second coder. Each coder kept a tally of mistakes made in coding by other coders; these tallies were monitored by the scoring supervisor, who apprised coders if they were getting careless or were consistently miscoding a particular question. For the industry and occupation questions, 100 percent of the questionnaires were recoded by a second coder. To prevent the second coder from being influenced by the first person's codes, the first person masked the codes in all questionnaires that he or she coded. The questionnaires were passed on and coded by a second person. The first person's codes were then unmasked; if there was a discrepancy between two codes, one of the designated coders met with the persons involved to resolve the discrepancy and to assign the most accurate code to the response.

7.2.2 Scoring Simulation Tasks

As the first shipments of booklets were received at ETS, copies were made of actual responses to the tasks. Staff members, including the test developer and scoring supervisor, scored these sample responses, using either the scoring guides developed for the young adult tasks or guides prepared during the development of the new tasks. As staff scored the sample responses, they made adjustments to the scoring guides for the new tasks to reflect the kinds of responses people were making.

The sample papers comprised the training sets used to train the readers who would score the survey booklets. The purposes of the training were to familiarize the readers with the scoring guides and to ensure a high level of agreement among the readers. The scoring supervisor conducted the training of the readers as a group. She explained each task and its scoring guide and discussed sample responses that were representative of the score points in the guide. The readers then scored an additional 10 to 30 responses. The supervisor then went over the scores for the responses and discussed those that presented particular problems for the group. After the group training was completed, each reader scored all the tasks in about 140 booklets to give them practice in scoring actual booklets, as well as to give them an opportunity to score a common set of responses as a means of checking their grasp of all the scoring guides. A follow-up session was then held to discuss responses for which there was disagreement among the readers. The entire training process was completed in about three weeks.

Twenty percent of all the booklets were subject to a reader reliability check, which entailed a second reader scoring the responses in those books. To prevent the second reader from being influenced by the first reader's scores, the first reader masked the scores in every fifth booklet that he or she scored. These booklets were passed on and scored by a second reader. The first reader's scores were then unmasked; if there was a discrepancy between two scores, the scoring supervisor reviewed the response and discussed it with the readers involved.

The statistic used to report inter-reader reliability is the percentage of exact agreement—that is, the percentage of times two readers agreed exactly in their scores. As can be seen from Table 7-1, the data show a high degree of inter-reader reliability across all the tasks in the survey, ranging from a low of 88 percent to a high of 100 percent, with an average percent agreement of 97. For 133 out of 168 open-ended tasks or parts of tasks, the agreement was above 95 percent.

The inter-reader reliability for the 1992 survey compares very favorably with the reliability for the 1985 young adult literacy assessment, which arrived at inter-reader reliability in the same way. For the young adult survey, the percent of exact agreement ranged from a low of 86 to a high of 100. For 54 out of the 66 open-ended tasks that were scored, the agreement was 95 percent or above, and the average agreement across all items was 96 percent.

Table 7-1. Summary of inter-reader reliability

Survey	Lowest percent	Highest percent	Average percent across all tasks	Number of tasks with 95 percent or above
Adults in 1992	88	100	97	133 out of 168
Young adults in 1985	86	100	96	54 out of 66

7.3 Data Entry

The background questionnaire was designed to be read by a computerized scanning device. For most questions, field personnel filled in the oval next to the respondent's answer. For open-ended items, ETS staff translated the responses and filled in the ovals before shipping the documents to the scanning department. Responses on the screener were transferred to scannable documents by ETS personnel when the check-in process was complete, and the screener documents were batched and sent to the scanning department at regular intervals. For the exercise booklets, ETS staff transcribed the scores for the simulation tasks and the responses to multiple-choice items onto scannable answer sheets, which were then forwarded to the scanning department. The scanned data from screeners, background questionnaires, and assessment booklets were transmitted to magnetic tape, which was then sent to the ETS computer center. As each of the different instruments was processed, the data were transferred to a database on the main computer for editing.

7.4 Editing and Quality Control

Editing included an assessment of the internal logic and consistency of the data received. For example, data were examined for nonexistent housing locations or booklets, illogical or inconsistent responses, and multiple responses. Where indicated, an error listing was generated and sent back to the processing area, where the original document was retrieved and the discrepancies were corrected. If it was not possible to resolve a conflict in the data, the information was left in the form in which it was received. Where possible, however, conflicts were resolved. For example, in the infrequent cases in which field personnel provided more than one response to a single-response non-cognitive item, specific guidelines were developed to incorporate these responses consistently and accurately. The background questionnaires were also checked to make sure that the skip patterns had been followed, and all data errors were resolved. In addition, a random set of booklets was selected to provide an additional check on the accuracy of transferring information from booklets and answer sheets to the database.

Chapter 8

ESTIMATING LITERACY PROFICIENCIES WITH AND WITHOUT COGNITIVE DATA

Kentaro Yamamoto, Educational Testing Service

In any educational, social or political opinion survey, missing responses are always present. There can be many reasons why sampled individuals do not respond to a survey. Sometimes missing data can be ignored when tabulating and reporting survey results. If the reasons the data are missing are related to the outcome of the study, however, the missing responses will bias the results unless some adjustment can be made to counter the bias. Experience with other surveys and field test evidence from the National Adult Literacy Survey indicated that adults with lower levels of literacy were more likely than adults with higher proficiencies either to decline to respond to the survey at all, or to begin the assessment but not to complete it. Ignoring the pattern of missing data would have resulted in overestimating the literacy skills of adults in the United States.

Missing survey data. The target sample for the survey included 24,827 nationally representative housing units, of which 3,164 were vacant. Approximately 11.5 percent of the households that were occupied at the time of data collection refused to participate in the survey, and no detailed background information is available on this group. Of the households that agreed to participate in the study, the interviewers began by using a series of screening questions to obtain an accurate count of the number of age-eligible persons in the household. Depending on the number of adults in the household, one or more persons were selected to participate in NALS. Respondents who did not answer a sufficient number of background questions were considered to be incomplete cases. Cases that were mostly incomplete could not be analyzed and were never incorporated into the database. Such cases were dealt with through weighting class adjustments for instrument nonresponse. (See Chapter 3.)

Missing cognitive data. The 26,091 persons, including prisoners, who agreed to respond to the survey answered extensive background questions during the interview about their age, country of birth, language(s) spoken or read, highest level of education completed, current educational aspirations, labor market status, current occupation and wages, voting behaviors, and reading habits. After answering the background questions, respondents were asked to complete the literacy tasks in the exercise booklet. Very easy tasks were placed first to encourage respondents to continue. Nevertheless, 1,364 (5 percent) of these respondents did not complete any cognitive tasks, and 1,630 (6 percent) responded to fewer than five tasks on at least one scale. For individuals who refused to continue after answering the background questions, no information is available about their performance on the cognitive tasks. Omitting these individuals from the analyses would have resulted in overestimates of the literacy skills of the national population as a whole and particularly of certain subpopulations. Special procedures were developed to estimate their literacy proficiencies. The first section of this chapter describes the method chosen to deal with the problem of missing cognitive data. The remainder discusses the results of related analyses.

8.1 THE NORMAL TREATMENT OF MISSING COGNITIVE DATA

In population assessments, unlike individual assessments, a matrix design for item sampling in which examinees respond to different subsets of cognitive questions is used to limit the burden on respondents. Due to the matrix design of the National Adult Literacy Survey, each respondent received only a fraction (three-thirteenths) of the literacy tasks. For every respondent, most of the tasks were not presented and could be considered missing, but this type of missing data was intentional. The tasks that were presented are sufficient to estimate parameters of the aggregate latent proficiencies distributions for subpopulations using item response theory models, even with a small number of completed responses

8.1.1 Omitted answers and questions not reached

For the literacy tasks that were presented, the missing responses occur in two distinct patterns: the respondent skipped over a question and responded to a subsequent question, or the respondent broke off the exercise booklet and did not attempt to respond to any subsequent questions. The two types of nonresponse are called "omitted" and "not reached" tasks, based on the pattern of responses to all the questions in each block.

- Omitted. In some cases, respondents skipped over a particular task, but attempted or completed one or more tasks that followed. This kind of missing response by definition, cannot be found at the end of a block.
- Not reached. In other cases, respondents spent all their time responding to preceding tasks, and did not reach tasks that appear later in a block. Tasks that were not attempted are found consecutively at the end of the blocks and are also termed "consecutively missing responses."

In the omitted response situation, there is a logical basis for assigning a wrong answer to a missing response. For the National Adult Literacy Survey, omitted cognitive responses were treated as wrong answers, on the assumption that respondents decided to skip them because they found these tasks too difficult. The treatment of an omitted response as a wrong answer is a logical imputation based on the circumstances that surround the missing data.

In contrast, 'not reached' cognitive responses were not assigned wrong answers. The assumption here is that respondents did not make a task-specific decision whether or not to respond to tasks that were not reached. Since there was not a sufficient logical basis for assigning a wrong answer, these responses remained missing data. Moreover, it is unlikely that 'not attempted' tasks occur because of speededness—that is, because the respondent was not given enough time to answer them. The assessment booklet was not rushed and is not a speeded test, but there were some practical time limits so that the interviewer would not have to spend an unreasonable number of hours collecting information. Logical imputations were used for missing responses to individual test questions. Statistical imputations were used to estimate unobserved proficiencies based on data that were incomplete by design.

8.1.2 Statistical imputation through scaling

Many analytic steps had to be taken to convert the information contained in responses to the 165 individual literacy tasks contained in the National Adult Literacy Survey Exercise Booklets into summary literacy estimates for populations (see Chapter 9). The first step used item response theory to estimate the relationships between assessment tasks and the underlying literacy skills that they are designed to measure. These relationships are quantified in the item parameters estimates available for each literacy task used in the survey.

In the next step, these relationships were used to build a likelihood function that measures how likely it is that a particular respondent has each value in a range of possible literacy skills. Many different levels of ability could generate a particular pattern of right or wrong answers, and some of these values are more likely than others. The likelihood function provides a measure of the probability that any given literacy proficiency underlies an individual's performance on the set of tasks that that individual completed. Achievement tests geared to measure individual abilities, such as the SAT and GED, generally take many hours to administer and obtain large numbers of tasks from each examinee. In these situations, the likelihood function is narrow and peaked, providing a good estimate of individual proficiency. However, the National Adult Literacy Survey kept the administration time to 45 minutes. The number of tasks taken by respondents on each of the prose, document, and quantitative literacy scales was small (a maximum of 12 to 15, not counting any missing cognitive responses). As a result, the likelihood functions tend at best to be broad and flat, providing poor estimates of individual proficiency. When responses were missing, the assessment offers even less information.

Using an approach called "marginal maximum likelihood," the next step estimates the relationship between background variables and underlying ability (see Mislevy, 1984, for a description of this method). The background data are summarized by more than 200 principal components, which capture 99 percent of the variance of the original background variables.

The results from the marginal maximum likelihood analysis are used to generate plausible values. Plausible values are imputed test scores that have approximately the same distribution (subject to some assumptions) as the unobserved underlying trait and are similar to imputed test scores. The plausible values are similar to regression imputations in which a random error from the appropriate distribution is added to the regression prediction. However, the plausible value method uses a Bayesian approach that combines an estimate of individual score likelihood with the regression imputation. In essence, this approach is a way to keep some of the variance in the original regression residual associated with the observation that contributed to it. That way, some biases may be reduced if the residual reflected variance due to background variables were excluded from the model. At the same time, this ensures that the mean and variance of the aggregate distribution (conditional on the background variables included in the model) is the same as regression imputation would yield.

In the case of the National Adult Literacy Survey, when the number of cognitive responses fell below five on a particular scale, the likelihood function became too wide and flat to consider the responses useful in estimating proficiency distributions. For the purposes of this discussion, any respondent who completed fewer than five tasks on any scale is considered to be an incomplete case.

Instead of a single point estimate of a respondent's proficiency, multiple (usually five) values are randomly sampled from the posterior distribution of each respondent and used as equivalent estimates of literacy skills for analytic purposes. These random draws from the posterior distribution are called plausible values. The essential idea of plausible values methodology is to represent what the true proficiency for an individual might have been, had it been observed. The methodology uses a small number of random draws from an empirically derived distribution of proficiency values that is conditional on the observed values of the assessment items and on background variables for each sampled adult. The random draws from the distribution can be considered to be representative values from the distribution of scale scores for all adults in the population with similar characteristics and identical patterns of item responses. The several draws from the distribution are different from one another in a way that quantifies the degree of precision with which the underlying conditioning model is estimated. The plausible value approach is described in Mislevy (1991).

These imputations are called "plausible values" because they are representative (and hence plausible) of the score distribution in the population of people who share the background characteristics of the individual with whom the plausible value is associated in the data. The more general term "imputed value" is not used, because this approach is a more specific kind of imputation that is different enough to deserve its own name. We note that plausible values are an extension of Rubin's (1987) multiple imputations. The imputations are designed to reproduce the aggregate relationships, not the individual scores. Detailed treatment of methodology and its application in the National Assessment of Educational Progress (NAEP) can be found in Mislevy, Johnson, and Muraki (1993) and Mislevy, Beaton, Kaplan, and Sheehan (1993).

8.2 REASONS COGNITIVE DATA WERE MISSING/NOT REACHED

Whenever a respondent broke off the National Adult Literacy Survey literacy exercises before the interviewer was able to lead the respondent through all of the tasks in the exercise booklet, the interviewer filled out the Non-interview Report (NIR) form (Exhibit 8-1). The Non-interview Report form asked how far the respondent got through the interview before it ended (Q.1), why the respondent was unable to begin or complete the interview (Q.2), if language was the problem, what language the respondent used (Q.3), and, if disability was the problem, the nature of the respondent's disability (Q.4) (See Appendix E.) The interviewers were not qualified to diagnose disabilities; such information had to be obtained from the respondent. The interviewer also recorded details about the nature of the reading and/or writing problem (Q.5) or the refusal (Q.6-7). The interviewer's field manual provided instructions for how to interpret the response options (Exhibit 8-2).

Exhibit 8-1. Non-interview Report form

BACKGROUND QUESTIONNAIRE
AND EXERCISE NON-INTERVIEW REPORT (NIR)

NAME OF INTERVIEWER: ______________________________ DATE: ________________

CAREFULLY PLACE LABEL INSIDE THIS BOX

⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	
①	①	①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②	②	②
③	③	③	③	③	③	③	③	③	③	
④	④	④	④	④	④	④	④	④	④	
⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	
⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	
⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	
⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	
⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	

ENTER BOOKLET NUMBER

⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①
②	②	②	②	②	②
③	③	③	③	③	③
④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨

1. What is the status of this case?

BQ nonresponse: respondent did not begin BQ .... ①
BQ nonresponse: respondent started but did not complete BQ .... ②
EX nonresponse: respondent completed BQ but did not start Exercise .... ③
EX nonresponse: respondent completed BQ and started but did not complete Exercise (Enter booklet ID number above) .... ④

2. Why were you unable to begin or complete the Background Questionnaire or Exercise?

Non-English language .... ① **(Q3)**
Physical or mental disability .... ② **(Q4)**
Reading and/or writing difficulty .... ③ **(Q5)**
Respondent refused .... ④ **(Q6)**
Someone refused for respondent .... ⑤ **(Q6)**
Maximum calls .... ⑥ **(Q7)**
Unavailable during field period .... ⑦ **(Q7)**
Other reason .... ⑧ **(Q7)**

3. Which non-English language did the respondent speak, read, and/or write?

Spanish .... ① **(BOX 1)**
Other language (SPECIFY): ____________________ .... ② } **(END OF NIR)**
Could not determine .... ③ } **(END OF NIR)**

BOX 1. If this is a BQ nonresponse (Q.1 coded 1 or 2), go to Q.8. Otherwise, end of NIR.

4. Code the nature of respondent's disability. Code all that apply.

Learning disability .... ①
Any mental or emotional condition .... ②
Mental retardation .... ③
Hearing impairment .... ④
Blind/Visual impairment .... ⑤
Speech/Language impairment .... ⑥
A physical disability .... ⑦
Other (SPECIFY): ____________________ ⑧

BOX 2. If this is a BQ nonresponse (Q.1 coded 1 or 2), go to Q.8. Otherwise, end of NIR.

Exhibit 8-1. Non-interview Report form — Continued

5. Explain in detail the nature of the respondent's reading and/or writing problem.

____

____

____

END OF NIR

6. Explain the circumstances surrounding the refusal or breakoff, recording what the respondent (or individual refusing for the respondent) said. .

____

____

____

GO TO Q.8

7. Explain in detail why your were unable to begin or complete the Background Questionnaire or Exercise.

____

____

____

8. Record any information below that might help another interviewer complete the Background Questionnaire and/or the Exercise Booklet with this respondent.

____

____

____

Exhibit 8-2. Interviewer instructions for Non-interview Report form (NIR)

Q.2 Your answer to this question should be consistent with the final result code you assigned at either the Background Questionnaire level or the Exercise level as the reason for the non-response. Refer to the following for definitions of the answer categories.

1 *Non-English language* - The respondent was unable to speak or read English well enough to begin or to complete the Background Questionnaire or Exercise. (If the respondent's non-English language is Spanish, you should have made an effort to use an adult to assist with the Spanish version of the Background Questionnaire.)

In determining when to assign this code, remember that there may be some respondents who cannot speak or understand English when spoken to but who are able to read in English. If a R is unable to speak or understand English, our rules state that he/she is unable to participate in the Background Questionnaire interview. However, if they can read English, they can complete the Exercise. If a non-English speaking respondent is able to complete the exercise, then try to use a translator to administer the BQ in the respondent's non-English language. This is the ONLY situation where a translator can be used to obtain a BQ interview in a language other than English or Spanish.

2 *Physical or mental disability* - The respondent was unable to complete or to begin the Background Questionnaire or Exercise due to a physical or mental impairment, such as a chronic illness, vision problems, hearing problems, or a psychological disorder.

3 *Reading and/or writing difficulty* - The respondent was unable to begin or to complete the Exercise due to insufficient reading and/or writing skills. This category should only include those respondents who cannot read or write for reasons other than a physical or mental disability.

4 *Respondent Refused* - The respondent was unwilling to begin or to complete the Background Questionnaire or Exercise.

5 *Someone refused for respondent* - An adult, not necessarily a household member, refused to allow the respondent to begin or to continue with the Background Questionnaire or Exercise. (Most commonly, this will be parents refusing for teenagers, or adults refusing for their elderly parents.)

6 *Maximum calls* - You were unable to complete the Background Questionnaire and Exercise after making at least four attempts (in addition to the Screener attempts) on different days and at different times.

7 *Unavailable during field period* - The selected respondent was out of town or was otherwise not available at anytime during the field period. This includes non-chronic illness that prevents participation for the duration of the field period.

8 *Other reason* - The respondent did not begin or complete the Background Questionnaire or Exercise for any other reason not covered elsewhere in this code structure.

Q.3 If you did not begin or complete the Background Questionnaire or Exercise because of a non-English language problem, that is, the respondent did not speak, read, or write English, code the respondent's non-English language at this question. (If the language is Spanish, you should have made an effort to use an adult to administer the Spanish version of the Background Questionnaire. Remember, however, a household translator cannot be used for any language other than Spanish.)

Q.4 If the respondent is unable to begin or to complete the Background Questionnaire or Exercise because of a physical or mental disability, we want to know more about the type of disability at Q.4. Some disabilities you will be able to observe readily. Other, like a learning disability or mental retardation, may be evident only if someone mentions the problem. It is also possible that a respondent could have more than one of the specified disabilities, so code all that apply. The following examples and definitions of some specific disabilities are to be used as guidelines to help you determine more accurately, the nature of the respondent's disability.

1 *Learning disability*: If a respondent says that he/she cannot participate in the exercise because of a learning disorder or learning disability, this code would apply. The most common type of learning disability is dyslexia, a type of impairment that affects a person's ability to read. If the respondent mentions dyslexia, or says that he/she has some other form of learning disability, this code would apply. Note that a learning disability should not prevent the respondent from participating in the Background Questionnaire interview. In addition, not all learning disorders would prevent a person from completing the exercise. For example, a person with Attention Deficit Disorder (ADD) is capable of reading and may have no difficulty with completing the Exercise. Persons reporting a learning disability should still be encouraged to attempt the tasks in the Exercise Booklet.

2 *Any mental or emotional condition*: This includes psychological disorders, dementia, autism, nervous breakdowns, severe depression, schizophrenia, etc.

3 *Mental retardation*: A person who is mentally slow or delayed in mental development. This can be caused by a birth defect, a congenital condition such as Down's Syndrome, or accident or injury to the brain that results in brain damage.

4 *Hearing impairment*: For R's who are totally (profoundly) deaf, this code should be used to describe a BQ nonresponse, since a totally deaf respondent cannot participate in the Background Questionnaire interview. R's who are partially deaf or hard of hearing, must be offered the opportunity to try the Exercise. If, however, they cannot begin or complete the Exercise because of a hearing impairment, use this code to describe the situation. Respondents with hearing impairments should be encouraged to use a hearing aid(s) or other device that would enable them to take part in the Background Questionnaire and Exercise.

5 *Blind/Visual impairment*: If R is totally blind, he/she can participate in the Background Questionnaire, but not in the Exercise, so this code would be used to describe an Exercise nonresponse. fl's who have a visual impairment must be offered the opportunity to try the Exercise. If, however, they cannot begin or complete the Exercise because of a visual impairment, use this code to describe the situation. Respondents with visual impairments should be encouraged to use glasses or other devices that would enable them to take part in the Exercise.

6 *Speech/Language impairment*: Includes conditions such as severe stuttering that disable speech communication in English other than lack of proficiency in English. This code also includes aphasia (acquired impairment of verbal behavior caused by brain damage that could impair the R's ability to use and/or understand words).

7 *A physical disability*: Includes physical conditions that interfere with an ability to perform the Exercise such as paralysis, amputation of hand/arm, lack of muscle control, etc.

8 *Other*: Includes Alzheimer's disease, senility and any other condition not covered elsewhere in this code structure.

For a respondent who has had a stroke or suffers from another medical condition that would prevent him/her from completing the Exercise, you should code the symptom(s) the respondent suffers from. For example, a stroke victim who is left paralyzed, blind, and unable to speak would be coded as "5 - Blind/visual impairment," "6 - Speech/Language impairment," and "7 - A physical disability."

Q 5 Respondents with reading and writing difficulties are encouraged to complete the Background Questionnaire and attempt the Exercise. Respondents who cannot read at all, or very little, may decide not to attempt the Exercise when you hand them the Exercise Booklet. If the respondent volunteers information on the extent of his or her reading ability at this time, record it here.

8.2.1 Non-interview Reports and low literacy skills

Table 8-1 shows the distribution of the nonresponses by reasons for not responding according to the Non-interview Report, in order from most to least frequently cited. Some of the reasons cited could reasonably be interpreted as indicating that the respondent had insufficient literacy skills to successfully complete the assessment, while other reasons did not provide any basis for inferring low literacy skills.

Non-English language. Some individuals (2.7 percent of the population) did not complete the assessment because they had difficulty reading in the English language. In most cases, this was because English was their second language. Nonresponse for these individuals was taken to indicate inability to read and write in English, but not necessarily indicating any inability in other languages.

Table 8-1. Distribution of nonresponses, by reasons for nonresponse

Number of cognitive tasks completed and reported reason (if fewer than five completed)	Percentage of adults	Presumed relation of reason to literacy skills
Five or more tasks	87.8	
Fewer than five tasks		
Non-English language	2.9	Related to literacy
Refused	2.7	Not related to literacy
Reading and/or writing difficulty	2.1	Related to literacy
Physical disability, including visual	1.9	Not related to literacy
Other, or unknown	1.8	Not related to literacy
Mental disability, including retardation, learning disability and other mental/ emotional condition	0.6	Related to literacy
Total	100	

Refusal. Some individuals (2.7 percent) refused to complete the cognitive tasks in the survey, despite efforts to keep the time burden to a minimum. A refusal could not reasonably be interpreted as indicating that the respondent had insufficient literacy skills to successfully complete the assessment, although adults with low literacy skills might well try to conceal their inability by refusing to be assessed.

Reading or writing difficulty. Some individuals could not be assessed because they had difficulty with reading or writing (2.1 percent). This category did not include those respondents who could not read or write due to a physical or mental disability, who were classified elsewhere. The assessment was not designed to investigate in detail the nature of these language difficulties; rather, the interviewers simply recorded the information. Use of this category was taken to indicate inability to read and write in English.

Physical disability. To answer the cognitive tasks, respondents had to be able to read materials and respond to tasks in writing without help. Individuals with physical disabilities (1.9 percent) such as visual impairment or lack of motor skills could not be assessed because they were physically unable to produce written responses. Their reason for not responding was considered unrelated to literacy skills because they may have been able to respond successfully if the tasks had been presented in a different mode (i.e., not paper and pencil).

Other. Some respondent (1.8 percent) did not respond to the literacy tasks for one of the following reasons: interviewer exceeded maximum number of contacts without making an appointment, respondent unavailable, other unspecified reasons, or missing information. The major reason for missing information was due to difficulty in determining why the respondent did not complete the survey. An inability to reach the potential respondent could not reasonably be interpreted as indicating that the respondent had insufficient literacy skills to successfully complete the assessment. Although some individuals with missing information might belong to one of following literacy-related categories, there was not enough information to make such a determination.

Mental or learning disability. Some adults (0.6 percent) did not respond to the survey because they (or others who knew them) identified themselves as having a mental/emotional condition, mental retardation, or a learning disability. In combination with missing cognitive data, use of this category was taken to indicate inability to read and write in English. Some adults may have had less severe versions of these conditions and responded to the survey. Only those who did not complete the cognitive portion of the survey were considered unable to read and write in English.

In studies of the mentally disabled population, about 1 to 2 percent of the total population are reported to be mentally disabled. Such disabilities may be present at birth or may be related to other factors, such as aging. The National Adult Literacy Survey included every household member within a specified age range, and consequently included more of the mentally disabled than typically found in surveys of children enrolled in educational institutions. For example, mentally disabled or learning disabled populations have not in the past been fully included in the sampling frameworks of traditional school-based assessments such as the National Assessment of Educational Progress.

The specific reasons for not responding can be summarized as related to literacy (i.e., the individual was unable to respond) or unrelated to literacy (i.e., the individual was unwilling to respond or did not provide a reason for not responding). Literacy-related reasons included difficulty with reading or writing, inability to read or write in English, or mental or learning disabilities. Unrelated reasons included interruptions, time conflicts, or physical disabilities (such as orthopedic, visual, or hearing impairments. When the reasons for stopping were unknown, there was no basis for inferring low literacy skills.

8.2.2 Internal evidence for the validity of reasons

If the reported reasons for not providing data were valid, then the reports could be used to improve the treatment of missing responses. While no follow-up interviews were possible, some internal validity checks are possible. The National Adult Literacy Survey collected extensive background information, some of which can be shown to be related to the reasons cited for nonresponse. Table 8-2 displays the distribution of the reasons respondents cited for providing insufficient cognitive data by age, language spoken when growing up, race/ethnicity, and educational attainment. Standard errors for Table 8-2 are provided in Appendix Q.

Age and reasons. Given the infirmities that accumulate as people grow older, one might expect older adults to be more likely than younger adults to state physical disability as a reason for nonresponse. Similarly, experience with many surveys have shown that older adults are more likely to refuse to participate in any kind of survey than younger adults. Table 8-2 shows that adults 65 or over, when they provide insufficient cognitive data, are more likely than adults under 65 to cite a physical disability and more likely not to explain their refusal than are younger adults. While adults 65 or older are more likely than younger adults to give reasons unrelated to literacy (18 percent compared to 4 percent), they are also more likely to cite reasons unrelated to literacy (11 percent compared to 5 percent). Most of this difference can be attributed to the greater frequency of reading or writing difficulties other than physical.

Race/ethnicity and reasons. Twenty-seven percent of the Hispanic population,17 percent of the Asian population, 8 percent of the black population, and 2 percent of the White population cited reasons related to literacy for nonresponse, most of which were problems with English. There were no marked differences among these groups in the proportion reporting reasons unrelated to literacy, such as refusal. It is likely, however, that much of the differences among the racial/ethnic groups are attributable to the language they learned when they were growing up. Immigrants are much more common among Hispanics and Asians than among blacks or Whites.

Native language and reasons. Individuals who spoke English at home when they were growing up tend to develop better English literacy skills than those who did not speak English early in their lives. Table 8-2 shows that adults who spoke English at home when they were growing up, when they provided insufficient cognitive data, almost never cited language as a reason, while about a third of those who had not spoken English at home indicated literacy-related reasons, and language-related reasons in particular, for their nonresponse.

Table 8-2. Percentage citing various reasons for providing insufficient cognitive data, by age, language spoken while growing up, race/ethnicity, and education: 1992 adults

Population group	Sufficient cognitive data present	Insufficient cognitive data present								Total sample size
		Reasons related to literacy				Reasons unrelated to literacy				
		Total	Non English language	Mental disability	Reading or writing difficulty	Total	Refused	Physical disability	Other, Unknown, No answer	
Total	88	6	3	1	2	6	3	2	2	26091
Age										
16 to 65	91	5	3	0	1	4	2	1	1	23877
65 and older	71	11	3	2	6	18	5	10	4	2214
Race/Ethnicity										
Hispanic	68	27	24	0	3	5	2	1	2	3126
Asian	78	17	15	1	1	5	2	2	2	390
Black	84	8	1	1	6	8	3	3	3	4963
White	91	2	0	1	1	6	3	2	2	17292
Other or missing	83	10	8	0	2	7	1	3	3	320
Language spoken while growing up										
English	91	2	0	1	2	7	3	2	2	23339
Languages other than English	60	34	29	1	5	6	2	2	2	2715
Missing data	92	1	0	0	1	7	0	0	7	37
English spoken while growing up										
Black	84	7	0	1	6	8	3	3	3	4881
Hispanic	93	2	1	0	1	4	2	1	2	1239
Asian	90	0	0	0	0	10	3	4	3	134
White	92	2	0	1	1	6	3	2	2	16824
English not spoken while growing up										
Black	56	40	38	2	0	5	3	0	2	69
Hispanic	52	43	38	0	5	5	2	1	2	1887
Asian	72	24	22	1	1	4	1	1	2	256
White	74	16	6	2	8	9	3	3	3	448
Education										
0 to 8 years	49	34	18	2	14	17	4	8	5	2167
9 to 12 years	83	7	3	1	3	10	4	3	3	3311
HS grad or GED	91	3	1	0	1	6	3	1	2	7169
Some postsecondary	95	1	1	0	0	4	2	1	1	7620
College grad	96	0	0	0	0	4	2	1	1	4787
Other or missing	93	5	4	1	1	2	1	0	0	1037
Total sample size	23097	1547	835	149	563	1447	614	386	447	

Native language, race/ethnicity, and reasons. When comparing the racial/ethnic distribution of reasons for providing insufficient cognitive data within groups who either spoke or did not speak English at home while growing up, the racial/ethnic disparities in reasons were quite different. Among those who had spoken English at home while growing up, virtually no black, Asian, Hispanic, or White adults reported language as their reason, so there were no longer any racial/ethnic differences. Among those who had spoken not English at home growing up, 38 percent of the Hispanic population, 22 percent of the Asian population, 38 percent of the (tiny) black population, and 6 percent of the White population cited problems with the English language as their reason for not providing sufficient cognitive data.

It should be noted that black adults were more likely than those in the other racial/ethnic groups to cite a reading or writing difficulty unrelated to English or a physical disability, and that this pattern did not occur for blacks who had not spoken English at home when growing up.

Education and reasons. Since literacy skills are learned in school, those with higher levels of education should be much less likely than those with lower levels to cite literacy-related reasons for providing insufficient cognitive data. Table 8-2 shows that this group of reasons was most prevalent among those who dropped out of school before high school (34 percent), and almost totally absent among those who attended at least some postsecondary education or more (1 percent or less). The lowest educated group not only had more problems with English than those who were more educated, they also had more reading and writing difficulties, and more physical and other reasons which were not directly related to literacy.

The above analysis provides some evidence that adults' self-reported reasons for providing insufficient cognitive data were internally consistent with their other relevant attributes, such as their age, educational attainment, and native language. It is likely that there were some errors in classifying the reasons that some adults failed to complete the assessment. Some adults may have given an explanation that reflected badly on their literacy skills simply because they found the task of completing the literacy assessment too burdensome. Perhaps they could have performed better had they tried harder. The assumption that such adults are unable to succeed with the literacy tasks may be too strong, and the assignment of wrong answers may underestimate their skills. Other adults may have anticipated "failure" in the assessment, yet concealed their lack of literacy skills by citing other reasons for not responding, or by refusing to explain their reason. The assumption that these adults are just like others in their demographic group may also be too strong, and the failure to assign wrong answers may overestimate their skills. The available data are insufficient to assess which kind of classification error occurred more often, but to some extent the errors would tend to counterbalance one another. The consistency of the findings lends some confidence in the validity and accuracy of the reported reasons for nonresponse.

8.3 USING 'REASONS' TO IMPROVE TREATMENT OF MISSING COGNITIVE DATA

The interviewers reported on the Non-interview Report form the respondent's reasons for nonresponse. The data concerning reasons for missing cognitive data provided the basis for making logical imputations of what the missing answers would have been had the respondent completed the exercise booklet. At the request of the National Center for Educational Statistics, ETS experimented with several ways of making logical imputations based on the Non-interview Report.

8.3.1 Five logical imputation methods considered

At the request of the National Center for Education Statistics, ETS used the National Adult Literacy Survey field test data to develop five logical imputation methods for using the information from the Non-interview Report form. The methods were only applied when the respondent answered fewer than five tasks on at least one scale. In addition, before the methods were applied, any omitted responses were assigned wrong answers, so that the logical imputations were applied only to missing tasks found consecutively at the end of the blocks. Since only missing data were logically imputed, no actual responses were ever overwritten by any of the five methods. Below, the implications of the resulting proficiency estimates are compared for various population subgroups.

The five methods differ in the extent to which they make use of the literacy-related reasons found in the Non-interview Report form and in the number of logical imputations carried out. Three of the five logical imputation methods made no use of the Non-interview Report data. The most drastic method assigned incorrect answers to *all* consecutively missing responses. A less drastic, but still uninformed method assigned incorrect answers to missing tasks among *only the first five* consecutively missing responses. Here the consecutively missing responses to the first five tasks were treated as incorrect responses, and the remaining portion of the consecutively missing responses were treated as not reached. The least informed method never imputed incorrect answers to any missing tasks, regardless of the information on the Non-interview Report.

Two of the five methods were informed by the reasons found in the Non-interview Report. The more drastic method of the two assigned incorrect answers to *all* consecutively missing responses of those whose reasons for nonresponse were literacy-related, and left alone the missing task responses of those indicating no reason, or a reason unrelated to literacy. The less drastic method assigned incorrect answers to *only the first five* consecutively missing responses of those whose reasons for nonresponse were literacy-related, and left alone all other missing task responses.

Table 8-3 summarizes the task assignments of the five methods, according to how the information from the Non-interview Report was used and the number of missing tasks that were imputed. The three methods that are un-informed by the reasons for nonresponse do not reflect reality as well as the others. Two of the methods assign wrong answers equally to those who refuse or have a physical disability as well as to those who have reading or writing or language difficulty. This is inconsistent with the presumed relationship of the reasons given to low literacy skills described in the previous section. Another method never assigns wrong answers, so it ignores reality in a different way. This method may not be so extreme if the survey population were fairly uniform—for example, an in-school population. The adult population as a whole is not uniform, however. For example, this population includes non-English speakers who reside in the United States, some of whom do not read English at all.

Table 8-3. Methods for the treatment of consecutively missing tasks, by degree of imputation and use of information from Non-interview Report

Reasons for Missing		Treatment of Missing	Method
No use of reasons		All missing changed to wrong	I
		Only missing among the first 5 responses changed to wrong; missing treated as not reached if in items 5 or greater	IV
		Missing items treated as not reached	V
Use of reasons	Related to literacy	Change all missing to wrong	II
		If missing in the first 5 items change missing to wrong; remaining missing items are treated as not reached	III
	Not related to literacy	Missing treated as not reached	II, III

The methods that did assign wrong answers varied in the number of wrong answer assignments made. A greater degree of logical imputation was provided by assigning wrong answers to all missing responses, while a lesser impact was provided by limiting the assignment of wrong answers to only the missing responses among the first five tasks on each scale. The likelihood functions derived from the latter are less informative than those derived from the former.

Each of the five methods has an impact on the likelihood function of the respondents, and hence on the posterior distribution from which population estimates are computed. Regardless of which of the five methods is used, the prior distribution is the same, since the relationship between literacy skills and background characteristics is the same for everyone. However, the posterior distribution could be different, since it is the product of the prior distribution and the likelihood function.

A respondent's plausible values are drawn from a posterior distribution that is the multiple of two functions, the conditional distribution of proficiency, given the pattern of background variables, and the likelihood function of proficiency, given the pattern of responses to the cognitive tasks. The conditional distribution of proficiency, given the pattern of background variables was not considered to be different for responding and nonresponding individuals. With this assumption, missing cognitive data that were not assigned wrong answers have no bearing on the proficiency estimates, given identical background information. In other words, the higher rate of refusals within certain subpopulations would not lower the mean proficiencies of those subpopulations. Any subpopulation's heterogeneity is evidenced by the wide distribution of background variables. Respondents that provided insufficient cognitive information may be differentially distributed within a subpopulation. In such cases, the proficiency distribution for nonrespondents would differ from the distribution of respondents due to differences in background variables.

8.3.2 Five methods applied to 1991 field test data

The field test was designed to study the effectiveness of a variety of procedures planned for the full-scale survey. The field test was conducted in a sample of 16 primary sampling units (PSUs), consisting of counties or groups of counties representing the 48 contiguous states. The PSUs were selected based on several key variables, including region of the country, urbanicity, race/ethnicity, and the average income/education level of persons residing in the selected PSUs. On average, 21 segments (consisting of census blocks) were selected within each PSU, with about eight households selected in each segment. Thirty-eight interviewers carried out the survey, in most cases two per PSU.

The survey instruments included a screener, a background questionnaire, and an exercise booklet. The screener enumerated household members in order to select an eligible respondent. The background questionnaire collected information in six areas: demographic data, language background, education, political and social participation, labor force participation, and literacy activities. The exercise booklet consisted of three 15-minute sections of prose, document, and quantitative tasks. The field test design included the administration of about 100 new literacy tasks. Interviewers were instructed to introduce the incentive after the household composition had been determined and the eligible respondent(s) selected. Respondents who completed the background questionnaire and agreed to complete the exercise booklet were given an incentive check after completing or attempting to complete the exercise booklet.

Of the 2,774 households in the sample, 12 percent were either vacant or did not satisfy the definitions of a dwelling unit at the time of screening. Among the 2,438 eligible households, 88 percent completed the screener, 6 percent refused to participate in the study, and 5 percent did not complete the screener for other reasons. From the households completing the screener, 2,288 eligible respondents were selected for the background questionnaire and the exercise booklet. Of these, 412 did not complete the background questionnaire; 63 percent were refusals, 10 percent had some type of mental or physical disability, and the remainder (27 percent) were nonrespondents for other reasons.

The completed task responses were scored right or wrong (or in some cases assigned wrong for omits or identified as missing for not reached); the five methods for assigning wrong answers to consecutively missing cognitive data were applied; and item response theory scaling procedures (Chapter 9) were used to to estimate provisional literacy proficiency scores. Unlike the scores of the full-scale survey, the literacy scales defined for the field test (prose, document, and quantitative) ranged from 0 to 100, with a mean of 50 and a standard deviation of 10. Using a different scale here helps us to see the implications of alternatives for treating missing responses.

Table 8-4 presents sample sizes and prose literacy scale results for those who responded to more than five prose tasks and those who responded to fewer than five prose tasks. Table 8-5 presents similar results for the total of the two groups. Standard errors are not presented because field test results cannot be considered population estimates. Furthermore, comparisons of different alternatives for treating missing responses to cognitive items are not influenced by sampling variation. There is only one field test sample. The observed differences in average prose literacy across columns in Table 8.4 result from variation in the number of incorrect answers imputed by the five methods. The observed difference between rows result from differences in the number of blank responses from various types of responses.

It was expected that including the nonresponding individuals with individuals who responded to more than five tasks per scale would result in somewhat lower overall proficiency means. The top line of Table 8-4 shows that this was indeed the case. The average prose proficiency of all those who completed at least five tasks on each scale was 50, compared to scores of 29 to 43 (depending on the assignment method) for those for whom fewer than five cognitive tasks were available. The corresponding line in Table 8-5 indicates that the combined average of the two groups was 48 to 50 (again depending on the assignment method). The minimal impact comes about because the percentage of respondents with insufficient cognitive data is greatly outweighed by those with sufficient data in the combined score. Among those with insufficient cognitive data, the difference between the assignment method with the least impact and that with the most was 14 points (29 compared to 43).

Since gender is not related to reasons for nonresponse, mean proficiencies of men and women are affected by all five assignment methods nearly equally, both within the group with fewer than five cognitive tasks available (Table 8-4) and in the combined total (Table 8-5). In addition, the prose scores of both men and women who provided sufficient cognitive information (50) were much higher than those who provided insufficient data (29 to 44). Among those with insufficient cognitive data, the difference between the assignment method with the least impact and that with the most was 15 points for men (29 compared to 44) and 14 points for women (29 compared to 43).

However, some population groups differ in the distribution of their reasons for nonresponse in ways that could produce different results depending on which of the five methods is used. Educational level interacted with proficiency means for all five methods. Among those with 0 to 8 years of education, the prose scores of those who provided sufficient cognitive information (38) were higher than those who provided insufficient data (27 to 32), except for the method with no logical imputations (38). With education, the differences among the methods showed distinct patterns. Among those who provided sufficient cognitive information, those with the least schooling had scores 19 points lower than those with the most (38 for 0 to 8 years, compared to 57 for 2 or 4 year degree). This difference was reproduced within the two methods for assigning wrong answers that took into account the reasons for nonresponse (corresponding differences were 18 and 19 points for the 'all missing' and the 'first five missing' methods of wrong answers for literary related reasons) (Table 8-4). However, this difference was smaller for the methods that did not take reasons into account (6 and 9 points for the 'all missing' and the 'first five missing' methods). This difference was 16 points for the method that did not involve logical imputations. Among those with insufficient cognitive data, the difference between the assignment method with the least impact and that with the most was 11 points for those with 0 to 8 years of schooling (27 compared to 38) and 19 points for those with a 2 or 4 year degree (33 compared to 54).

Income level also interacted with proficiency means for all five methods. Among respondents with household income from 5 to 10 thousand dollars per year, the prose scores of those who provided sufficient cognitive information (45) were higher than those who provided insufficient data (28 to 40) (Table 8-4), depending on the method. Like educational background, the income level differences among the methods showed distinct patterns. Among those who provided sufficient cognitive information, those with very low household income—from 5 to 10 thousand dollars per year—had scores 10 points lower than those with the most—50 thousand or more dollars per year (45 for $5-9,999, compared to 55 for $50,000 or more). This difference was reproduced within the two methods for assigning wrong answers that took into account the reasons for nonresponse (corresponding differences were 11 and 12 points for the 'all missing' and the 'first five missing' methods). However, this difference was smaller for the methods that did not take reasons into account (5 and 7 points for the 'all missing' and the 'first five missing' methods). This difference was 12 points for the method that did not involve logical imputations. Among those with insufficient cognitive data, the difference between the assignment method with the least impact and that

Table 8-4. Average prose proficiencies of adults with and without at least five cognitive tasks on each literacy scale, after assignment of wrong answers to missing cognitive tasks, by method of assignment and population group

Population group	At least 5 cognitive tasks available for each scale		Fewer than 5 cognitive tasks available for at least one scale					
	Field test sample size	Average prose proficiency	Field test sample size	Average prose proficiency				
				Wrong answers assigned for both types of reasons		Wrong answers for literacy-related reasons only		For neither type of reason
		No wrong answers assigned		To all missing	Only to first 5 missing	To all missing	Only to first 5 missing	No wrong answers assigned
Total	1707	50	154	29	33	38	39	43
Sex								
Male	775	50	63	29	33	38	40	44
Female	917	50	88	29	32	37	39	43
Race/Ethnicity								
White	1370	51	95	30	34	42	43	45
Black	171	42	18	27	29	35	35	38
Hispanic	124	46	32	28	31	30	32	40
Asian	26	50	8	-	-	-	-	-
Education								
Still in high school	68	48	1	-	-	-	-	-
0 to 8 years	94	38	49	27	29	30	32	38
9 to 12 years	180	42	25	28	31	37	38	40
GED	54	46	1	-	-	-	-	-
High school diploma	405	48	30	30	34	45	45	46
Some postsecondary	438	52	18	32	37	43	45	51
2 or 4 year degree	437	57	18	33	38	49	50	54
No education in U.S.	18	41	8	-	-	-	-	-
Age								
16 to 20	156	49	6	-	-	-	-	-
21 to 25	153	52	10	30	34	39	41	45
26 to 31	211	53	9	-	-	-	-	-
32 to 45	544	52	29	30	34	37	40	46
46 to 64	356	49	38	30	34	40	41	46
65 and older	287	44	62	28	31	38	39	41
Income								
<$5,000	61	46	7	-	-	-	-	-
$5,000-9,999	98	45	27	28	31	34	35	40
$10,000-14,999	142	46	17	28	30	34	36	40
$15,000-19,999	128	48	14	29	32	33	35	42
$20,000-29,999	241	49	17	28	32	39	40	42
$30,000-39,999	237	51	8	-	-	-	-	-
$40,000-49,999	204	51	9	-	-	-	-	-
$50,000+	465	55	16	33	38	45	47	52
Refused	40	46	24	31	35	46	46	47
Don't know	64	45	14	28	31	31	34	40

Table 8-5. Average prose proficiency scores of all adults after assignment of wrong answers to consecutively missing cognitive questions, on special field test 0-100 scale, by method of assignment and population group

Population group	Field test sample size	Average prose proficiency				
		Wrong answers assigned for both types of reasons		Wrong answers for literacy-related reasons only		For neither type of reason
		To all missing	Only to first 5 missing	To all missing	Only to first 5 missing	None
Total	1861	48	49	49	49	50
Sex						
Male	838	49	49	49	49	50
Female	1005	48	48	49	49	49
Race/Ethnicity						
White	1465	50	50	51	51	51
Black	189	41	41	42	42	42
Hispanic	156	42	43	43	43	45
Asian	34	45	46	45	46	48
Education						
Still in high school	69	48	48	48	48	48
0 to 8 years	143	34	35	35	36	38
9 to 12 years	205	41	41	42	42	42
GED	55	46	46	46	46	46
High school diploma	435	47	47	48	48	48
Some postsecondary	456	51	52	52	52	52
2 or 4 year degree	455	56	56	56	56	57
No education in U.S.	26	37	38	37	38	40
Age						
16 to 20	162	48	48	48	48	48
21 to 25	163	50	51	51	51	51
26 to 31	220	52	52	52	52	53
32 to 45	573	51	51	51	51	52
46 to 64	394	48	48	49	49	49
65 and older	349	41	42	43	43	44
Income						
<$5,000	68	44	45	45	45	46
$5,000-9,999	125	41	42	42	43	44
$10,000-14,999	159	44	44	44	45	45
$15,000-19,999	142	46	46	46	47	47
$20,000-29,999	258	48	48	48	48	48
$30,000-39,999	245	50	50	50	50	51
$40,000-49,999	213	50	50	51	51	51
$50,000+	481	54	54	54	54	55
Refused	64	40	42	46	46	46
Don't know	78	42	42	42	43	44

with the most was 13 points for those with 5 to 10 thousand in household income (27 compared to 40) and 18 points for those with 50 thousand or more dollars per year (33 compared to 52).

Not all minority populations were affected in the same way by each of the five methods. Among those who provided sufficient cognitive information, Whites had scores 9 points higher than blacks and 5 points higher than Hispanics (Table 8-4). The White-black difference was almost reproduced within the two methods for assigning wrong answers that took into account the reasons for nonresponse (corresponding differences were 7 and 8 for the 'all missing' and the 'first five missing' methods). The White-Hispanic difference was increased within the two methods for assigning wrong answers that took into account the reasons for nonresponse (corresponding differences were 12 and 11 for the 'all missing' and the 'first five missing' methods). However, the White-black difference was smaller for the methods that did not take reasons into account (3 and 5 points for the 'all missing' and the 'first five missing' methods). Similarly, the White-Hispanic difference was also smaller for the methods that did not take reasons into account (2 and 3 points for the 'all missing' and the 'first five missing' methods). For the method that did not involve logical imputations, the White-black difference was 7 points, and the White Hispanic difference was 5 points.

8.3.3 The method selected

Methods for assigning missing cognitive data without taking into account the reasons for non-response were found to weaken the educational, income, and racial/ethnic differences in literacy scores observed using field test data. Therefore, the methods that did not make use of the information from the Non-interview Report were ruled out. Among the two remaining methods, the cutoff at five tasks appeared arbitrary, and the consequences of assigning wrong answers to all of the missing tasks did not appear from the field test data to be unduly severe. After consultations with the Technical Review Committee and the Literacy Definition Committees advising the National Adult Literacy Survey project, a single method for assigning wrong answers was adopted for use in the full-scale survey, according to which all consecutively missing tasks for which the reason given was literacy related were assigned wrong answers and scored as if the respondent had attempted and failed with such tasks. This method was selected as the most viable approach for including individuals without cognitive data into the sample.

The field test data showed that for most groups, this procedure had little impact on the overall scores. The combined overall average prose literacy score dropped one point compared to the overall average of those with sufficient cognitive data (49 compared to 50). The same one-point drop occurred for Whites and for blacks, but the drop was magnified in ethnic groups for which use of English was not nearly universal. The Hispanic and Asian populations showed 3 point and 5 point drops respectively. This is because many of the nonresponding Hispanic and Asian persons indicated a literacy-related reason for not completing the assessment. Excluding the nonresponse samples would severely overestimate the literacy proficiencies of some subpopulations, namely the Hispanic and Asian populations because these groups contain a significant number of non-native speakers of English. To a lesser degree, the proficiency means of less educated, younger, and poorer subpopulations would also be overestimated.

All individuals without cognitive data who are included in the analyses have fairly extensive background information. Let us reiterate the method selected and its impact on the estimates of proficiency distributions for various subpopulations. This method imputes the proficiency scores of individuals who did not respond to at least five tasks. The procedure used background variables and self-reported reasons for nonresponse in addition to the functional relationship between background variables and proficiency scores for the total population. See Appendix N. for the *Impact of Treatment on Distribution of Scale Scores.*

8.4 Final Evaluation

In any survey of opinions, ability or demographic characteristics, missing responses are always present. The most commonly practiced and least desirable way to treat missing data is to ignore it. This practice assumes that missing cases are missing at random and that the remaining observed samples are representative of the target population. This practice would yield both biased and inaccurate proficiency distributions for some subpopulations if response rate are different among subpopulations, and consequently for the total population, as well.

The procedure utilized in the National Adult Literacy Survey classified nonrespondents into two separate groups. One group can be thought of as a very unable population, including those who do not use English, those with extreme reading difficulties, and those with some type of cognitive disability. The other group did not respond for reasons that are not strongly related to literacy in English; these include adults who simply refused to respond to the assessment tasks, as well as those with physical disabilities. Responses to the background variables indicate that those who did not respond to the cognitive items for the literacy-related reasons were disproportionately likely to be foreign born, to have less than a high school education, to be Hispanic or Asian/Pacific Islander, and to be age 65 or older. These variables are known to relate to English language proficiency and cognitive skills. Combined with other background information, there is strong evidence to support the notion that nonresponse to the cognitive items is not a random occurrence.

The above analysis assumes that adults' self-reported reasons for nonresponse are accurate and reliable. The accuracy of the Non-interview Report Form (NIR) information is particularly important because of its impact on the proficiency distributions, particularly for some subpopulations. It is highly unlikely that this level of consistency could have occurred if the NIR information were erroneous. In future assessments of this kind, however, it might be advantageous to incorporate a system for monitoring the reliability of the NIR data.

Chapter 9

SCALING AND PROFICIENCY ESTIMATES

Kentaro Yamamoto, Educational Testing Service

The National Adult Literacy Survey results are reported on the same three proficiency scales—prose, document, and quantitative—used for the NAEP 1985 young adult literacy assessment. This chapter describes the models and procedures used to scale the National Adult Literacy Survey results, to estimate respondents' proficiencies, and to conduct statistical analyses.

9.1 SCALING

The National Adult Literacy Survey gathered descriptive and proficiency information on 26,091 sampled respondents through a background questionnaire and a series of assessment booklets containing prose, document, and quantitative literacy tasks. Respondents were sampled using a four-stage stratified sampling method, as described in Chapter 2. In addition to the national sample, several other samples of respondents were surveyed using the same or similar instruments and mode of administration. Eleven states chose to participate in the concurrent State Adult Literacy Survey, each of which surveyed a sample of approximately 1,000 adults: California, Illinois, Indiana, Iowa, Louisiana, New Jersey, New York, Ohio, Pennsylvania, Texas, and Washington. Florida also surveyed approximately 1,000 adults, but at a later date. These supplementary samples allow results to be reported for these individual states; such information would not be possible if only the state's portion of the national sample were available for analysis. Another supplementary sample included 1,147 respondents incarcerated in 80 state and federal prisons.

All but 1,000 survey respondents who were living in households (that is, who were not in prison) received a monetary incentive of $20 for their participation. Previous studies on the use of incentive payments have found that the absence of an incentive lowers response rates, especially among respondents whose literacy proficiency is low. A response incentive payment of $15 was used in the 1985 young adult literacy assessment. For this reason, the payment of an incentive to the National Adult Literacy Survey respondents maintained comparability. At the request of the Office of Management and Budget, an experimental sample of 1,000 respondents did not receive any incentive, monetary or otherwise, in order to explore further the effects of incentives on the survey results. The results for this non-incentive sample were not included in the National Adult Literacy Survey reports, and are not included in this chapter.

Survey participants spent approximately 20 minutes answering a common set of background questions concerning their demographic characteristics, educational experiences, labor market experiences, and literacy-related activities. Responses to these background questions serve two major purposes. First, they provide a way to summarize the survey results using an array of descriptive variables, such as sex, age, educational attainment, and country of birth. Second, they increase the accuracy of the proficiency estimates for various subpopulations, as described later in this chapter.

The respondents spent the remainder of their time, approximately 45 minutes, completing a booklet of literacy tasks, measuring their prose, document, and quantitative skills. The assessment tasks administered in the National Adult Literacy Survey were created based on a definition of literacy drafted by a panel of experts in the field (see Chapter 4). Most of the cognitive tasks included in the assessment were open-ended or constructed-response questions that required respondents to provide a written answer. A small number of multiple-choice tasks were carried over from the earlier literacy surveys, making it possible to measure trends in performance and to compare the results from different assessments.

A large number of tasks had to be administered in the National Adult Literacy Survey to ensure that the survey would provide the broadest possible coverage of the literacy domains specified. Yet, no individual could be expected to respond to the entire set of 166 simulation tasks (tasks that simulate the demands that adults encounter when they interact with printed materials on a daily basis). Accordingly, the survey was designed using a variant of matrix sampling to give each participant a subset of the total pool of literacy tasks, while at the same time ensuring that each of the 166 tasks was administered to a nationally representative sample of adults.

Respondents' literacy proficiencies are estimated based on their performance on the cognitive tasks administered in the assessment. Unlike multiple-choice questions, which are commonly used in large-scale surveys, open-ended tasks such as those used in the National Adult Literacy Survey elicit a large variety of responses. Verbatim responses must be grouped in some way in order to summarize the performance results. Responses to the open-end tasks of the National Adult Literacy survey were classified into four categories: correct, incorrect, omitted, and not presented.

Since the National Adult Literacy Survey used a variant of matrix sampling and different respondents received different sets of tasks, it would be inappropriate to use any statistic based on the number of correct responses for reporting results, such as the proportion of tasks answered correctly. Differences in total scores (or statistics based on them) between respondents who took a different set of tasks may be caused by differences in respondents' abilities, differences in difficulty between the two sets of tasks, or both. Unless one makes very strong assumptions—for example, that the two sets of tasks are perfectly parallel—the performance of the two groups assessed in a matrix sampling arrangement cannot be directly compared using total score statistics. Moreover, task-by-task reporting ignores the similarities of subgroup comparisons that are common across tasks. Finally, using the average percentage of tasks answered correctly to estimate the proficiency means of examinees in a given subpopulation does not provide any other information about the distribution of skills within that subpopulation.

These limitations of conventional scoring methods can be overcome by using item response theory. When several tasks require similar skills, the response patterns should have some regularity. This regularity can be used to characterize both respondents and tasks in terms of a common scale, even when all respondents do not receive identical sets of tasks in their booklets. In this way, it becomes possible to discuss distributions of performance in a population, or subpopulation, and to estimate the relationships between proficiency and background variables.

The methods and procedures used to analyze the National Adult Literacy Survey results were carefully designed to capture most of the dominant data characteristics. Nevertheless, whatever procedure is used to aggregate data, a certain amount of information is lost when it does not fit the statistical model for proficiency estimates. The data that do not fit must be regarded as inessential to the analyses.

The design of the 1985 NAEP young adult literacy assessment established four proficiency domains—prose, document, quantitative, and reading. For the 1992 National Adult Literacy Survey, scaling was carried out separately for three of these four domains. The 1985 reading scale was dropped from the analyses because what the NAEP reading scale measures had changed in the intervening years. Use of the 1985 block of NAEP reading tasks would no longer be useful for comparisons to the 1992 NAEP reading assessment. The 1992 NAEP reading assessment had changed its block design to 25 minute reading blocks that would not fit the 15-minute block structure of the 1992 National Adult Literacy Survey. Accordingly, the three scales analyzed for the National Adult Literacy Survey were prose literacy, document literacy, and quantitative literacy, but not NAEP reading. By creating a separate scale for each of these domains, it remains possible to explore potential differences in subpopulation performance across these domains. Chapter 12 of this report discusses the rationale for using three distinct scales and examines the correlations among them.

9.2 SCALING METHODOLOGY

This section reviews the scaling model employed in the analyses of the National Adult Literacy Survey data and describes the plausible values methodology used for proficiency estimation.

9.2.1 The Scaling Model

The scaling model used for the National Adult Literacy Survey is the three-parameter logistic (3PL) model from item response theory (Birnbaum, 1968; Lord, 1980). It is a mathematical model for estimating the probability that a particular person will respond correctly to a particular task from a single domain of tasks. This probability is given as a function of a parameter characterizing the proficiency of a given person, and three parameters characterizing the properties of a given task. The following three-parameter logistic item response theory model was employed in the National Adult Literacy Survey:

$$P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + \frac{1 - c_i}{1 + e^{-1.7a_i(\theta_j - b_i)}} \qquad (1)$$

where

x_{ij}	is the response of person j to task i, 1 if correct and 0 if incorrect;
θ_j	is the proficiency of person j (note that a person with higher proficiency has a greater probability of responding correctly);
a_i	is the slope parameter of task i, characterizing its sensitivity to proficiency;
b_i	is its locator parameter, characterizing its difficulty; and
c_i	is its lower asymptote parameter estimated only for the multiple-choice tasks, reflecting possibly non-zero chances of correct response, even for persons with very low proficiencies; for open-ended tasks, c was fixed at zero.

Note that this is a monotonically increasing function with respect to θ; that is, the conditional probability of a correct response increases as the value of θ increases. In addition, a linear indeterminacy exists with respect to the values of θ_j, a_i, and b_i for a scale defined under the three-parameter model. In other words, for an arbitrary linear transformation of θ, say $\theta^* = M\theta + X$, the corresponding transformations $a^*_i = a_i/M$ and $b^*_i = Mb_i + X$ give:

$$P(x_{ij} = 1 | \theta^*_j, a^*_i, b^*_i, c^*_i) = P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i) \qquad (2)$$

Linear transformation of the scales was used to link the National Adult Literacy Survey scales to the 1985 young adult literacy assessment scales for gain purposes. The scale indeterminacy was resolved by setting an origin and unit size of θ to the reported scale means and standard deviations from 1985 young adult literacy assessment.

The main assumption of item response theory is conditional independence. In other words, item response probabilities depend only on θ (a measure of proficiency) and the specified item parameters, as opposed to depending on any demographic characteristics of examinees, or on any other items presented together in a test, or on the survey administration conditions. Controlling for θ, the probability of a correct response on one item is unrelated to the probability of a correct response on another given θ. This allows one to formulate the following joint probability of a particular response pattern x across a set of n items.

$$P(\mathbf{x}|\theta,\mathbf{a},\mathbf{b},\mathbf{c}) = \prod_{i=1}^{n} P_i(\theta)^{x_i}(1-P_i(\theta))^{1-x_i} \qquad (3)$$

By replacing the hypothetical response pattern with the real scored data, one can view the above function as a likelihood function that is to be maximized with a given set of item parameters. These item parameters were treated as known for the subsequent analyses.

Another assumption of the model is unidimensionality—that is, performance on a set of items is accounted for by a single variable. Although this assumption may be too strong, the use of the model is motivated by the need to summarize overall performance parsimoniously within a single domain. Hence, item parameters were estimated for each scale separately.

Testing the assumptions of the item response theory model, especially the assumption of conditional independence, is a critical part of the data analyses. Serious violation of the conditional independence assumption would undermine the accuracy and integrity of the results. Thus, while the item parameters were being estimated, empirical distribution of percentages correct conditional on θ and the item parameters were monitored across the adult sample of individuals 16 to 65 and the sample of adults over 65. For a few tasks, the percentages of correct responses obtained by the older sample were quite different from those obtained by the younger sample, and these tasks were dropped from the National Adult Literacy Survey analyses.

9.2.2 Design for Linking the 1992 Scales to the 1985 Scales

As previously noted, the prose, document, and quantitative literacy results for the National Adult Literacy Survey are reported on scales that were established in the 1985 young adult literacy assessment. Eighty-five (51 percent) of the tasks administered in the 1992 National Adult Literacy Survey were originally administered in 1985. The linkage between the scales from the two surveys is based on these tasks. In addition, 81 new tasks were developed for the National Adult Literacy Survey. A total of 166 tasks were administered in the 1992 survey. The composition of the National Adult Literacy Survey item pool is presented in Table 9-1.

Table 9-1. Composition of item pool for the National Adult Literacy Survey

Literacy scale	Number of tasks common to 1985 and 1992	Number of tasks in 1992 only	Total in 1992
Prose	14	27	41
Document	56	26	81
Quantitative	15	28	43
Total	85	81	166

A unidimensional item response theory model like the three-parameter logistic one employed in this study assumes that performance on all the items in a domain can, for the most part, be accounted for by a single (unobservable) proficiency variable. Subsequent linking and scaling analyses treated each scale separately—that is, a unique proficiency was assumed for each scale. As a result, the linking of corresponding scales was carried out for each scale separately. The three steps used to link the 1985 and 1992 scales are listed below.

1. Establish provisional item response theory scales through common item parameter calibrations based on a pooling of the 1992 and 1985 tasks.
2. Estimate the distribution of proficiencies on the provisional item response theory scales using plausible values.
3. Align the 1992 National Adult Literacy Survey scales to the 1985 scales by a linear transformation based on the common proficiency distribution of the 1985 sample.

9.2.3 Item Parameter Estimation

Identical item calibration procedures, described here in detail, were carried out separately for each of the three literacy scales. Using a modified version of Mislevy and Bock's (1982) BILOG computer program, the three-parameter logistic item response theory model was fit to each task (but with lower asymptote parameters fixed at zero for open-ended tasks) using sample weights.

The cognitive tasks administered in the 1985 young adult literacy assessment were used for several assessments and surveys, including the National Adult Literacy Survey, surveys in Oregon and Mississippi, the 1989-90 survey of job-seekers conducted for the U.S. Department of Labor, and a second Department of Labor assessment. In total, more than 40,000 individuals have responded to either the entire set or a subset of the 1985 young adult literacy assessment tasks. To obtain stable item parameter estimates and simplify scale linking procedures, the data accumulated from all surveys were included in a calibration sample. The current method of parameter calibration in effect puts all available survey results on a single provisional common scale. Only linear indeterminacy needed to be resolved in order to align the provisional scale to the reporting scale.

Sample weights were used during item calibration. It is known that different subpopulation distributions occur within different assessment samples. Such variations may arise because of differences in the characteristics of the target populations, the sampling design, or the randomness of sampling. For example, oversampling of racial/ethnic minority populations is often necessary to ensure a certain degree of accuracy in estimating group proficiencies. In such cases, the unweighted sample would not represent the targeted population correctly. Post-stratified weights take into account the sampling design, such as oversampling as well as the randomness of real data. By applying post-stratified weights, vital characteristics of the sample can be closely matched to the characteristics of the population. During calibration, the fit of item parameters is maximized in reference to the proficiency distribution of the calibration sample. When item parameters are being estimated, it is ideal to match the proficiency distribution of the calibration sample as closely as possible to that of the population. It is more critical when item calibration is done on the combined proficiency distribution of multiple assessment samples with great differences in proficiency distributions, such as the National Adult Literacy Survey. It was not as critical for the analysis of the 1985 young adult literacy assessment results because the young adult item parameters were estimated based on one sample.

To obtain unbiased parameter estimates, proficiency distributions for the separate assessment samples were estimated during calibration. In addition to the samples from the previous assessments, certain groups in the National Adult Literacy Survey respondents received separate proficiency distributions; those included adults age 16 to 64, those age 65 and older, prisoners, and respondents who received no monetary incentive for participating in the survey. It is known that the samples for each assessment came from somewhat different populations with different characteristics. In addition, the number of tasks administered varied in each assessment. The calibration procedure should take into account the possibility of systematic interaction of samples and tasks to generate unbiased estimates of sample distributions and item parameters. For that reason, a normal distribution with a unique mean and variance for each assessment population was estimated concurrently with item parameters. Estimated item parameters for each literacy scale are presented in Tables 9-2p, 9-2d, and 9-2q.

Model fit was evaluated at the task level by examining BILOG likelihood ratio chi-square statistics for each survey sample.[1] The fit was also evaluated by inspecting residuals from fitted item response curves. A typical plot is shown in Exhibit 9-1.

In Exhibit 9-1, the horizontal axis represents the provisional proficiency scale derived directly from the calibration procedure. The provisional scale is in standard units, without transformation to the 0 to 500 scale used for other purposes. The smooth curved line is the fitted three-parameter logistic item response curve. Each calibration sample is represented by a unique plot symbol. The five plot symbols represent the (approximate) expected proportions of correct responses at various points along the scale. The size of the plot symbols is proportional to the information available in the calibration data in that region of the scale. In general, the fit of the model was quite good. For some tasks, there was evidence that the estimated parameters did not fit certain assessment samples as well as other samples; however, this pattern was not consistently apparent for any one sample. Five tasks were dropped from calibration due to a lack of fit.

---

[1] The sampling distributions are probably not strictly χ^2 with the indicated degrees of freedom. Therefore, they were used as descriptive indices of relative model fit rather than as a statistical test of fit.

Table 9-2p. Prose literacy item descriptions and parameters for the National Adult Literacy Survey

Number	Description	A	SE(A)	B	SE(B)	C	SE(C)
AB21101	Swimmer: Underline sentence telling what Chanin ate	1.125	0.042	-1.901	0.048	0.000	0.000
AB21201	Swimmer: Age Chanin began to swim competitively	1.070	0.029	-1.124	0.027	0.000	0.000
AB30501	Technology: Underline sentence explaining action	0.590	0.015	0.593	0.022	0.000	0.000
AB30601	Technology: Orally explain info from article	0.915	0.023	0.347	0.018	0.000	0.000
AB31201	Dickinson: Describe what is expressed in poem	0.725	0.018	0.691	0.020	0.000	0.000
AB40901	Korean Jet: Give argument made in article	0.826	0.018	0.165	0.017	0.000	0.000
AB41001	Declaration: Describe what poem is about	0.622	0.020	-1.433	0.053	0.000	0.000
AB50101	Panel: Find information from article	0.466	0.016	2.112	0.057	0.000	0.000
AB50201	Panel: Determine surprising future headline	1.160	0.036	0.861	0.017	0.196	0.000
AB60201	Make out check: Write letter explaining bill error	1.240	0.027	-0.440	0.015	0.000	0.000
AB60601	Economic index: Underline sent. Explaining action	0.808	0.019	-0.319	0.021	0.000	0.000
AB70401	Almanac vitamins: List correct info from almanac	0.705	0.018	-0.765	0.029	0.000	0.000
AB71001	Instruction to return appliance: Indicate best note	1.377	0.042	-0.305	0.020	0.266	0.000
AB71101	Explain difference between 2 types of benefits	0.782	0.021	0.482	0.021	0.000	0.000
NC00301	"My Dream:" Find country in short story	0.892	0.034	-3.228	0.090	0.000	0.000
NC00401	"My Dream:" Underline sentence explaining action	0.765	0.016	-1.935	0.034	0.000	0.000
N010101	Marketing: List two facts	0.868	0.025	0.607	0.022	0.000	0.000
N010201	Marketing: Underline sentence explaining action	1.059	0.031	-0.402	0.022	0.000	0.000
N010301	Marketing: Give purpose of event	0.786	0.031	2.138	0.053	0.000	0.000
N080101	SSI: Mark correct information in article	1.328	0.051	-1.447	0.036	0.000	0.000
N080201	SSI: What must an SSI user accept if offered?	1.516	0.043	-0.389	0.017	0.000	0.000
N080301	SSI: What is most you can make to receive SSI?	0.618	0.021	0.486	0.028	0.000	0.000
N090601	Face off: What group will mandate safe cars?	1.878	0.064	-0.748	0.018	0.000	0.000
N090701	Face off: Find correct information in article	1.804	0.060	-0.699	0.018	0.000	0.000
N090801	Contrast views on fuel-efficiency vs. size of car	1.239	0.037	1.091	0.020	0.000	0.000
N100101	"Growing Up:" Find first buyer's name	1.466	0.052	-1.146	0.027	0.000	0.000
N100201	"Growing Up:" Determine correct day of delivery	1.297	0.037	-0.345	0.018	0.000	0.000
N100301	"Growing Up:" What reason given to stop selling?	1.187	0.034	-0.343	0.020	0.000	0.000
N100401	"Growing Up:" Compare approaches to selling mags	0.841	0.027	1.236	0.029	0.000	0.000
N110101	Blood pressure: Why difficult to know if high	0.988	0.032	-0.971	0.032	0.000	0.000
N110401	Jury: Length of time served by a juror	0.770	0.024	-0.191	0.027	0.000	0.000
N110501	Jury: Underline sentence explaining action	0.939	0.030	-0.730	0.030	0.000	0.000
N110601	Two challenges attorneys use to jurors	1.044	0.039	1.954	0.038	0.000	0.000
N120301	Ida Chen: What experience turned Ida toward law?	1.074	0.030	0.141	0.019	0.000	0.000
N120401	Two things Chen did to resolve discrimination conflicts	1.162	0.032	0.229	0.017	0.000	0.000
N120501	Ida Chen: Interpret phrase from article	0.926	0.037	2.107	0.048	0.000	0.000
N120901	Susan Butcher: Find number of wins of sled race	0.888	0.044	-2.061	0.080	0.000	0.000
N130201	Fueled: Determine phrase meaning	1.089	0.030	0.315	0.018	0.000	0.000
N130301	Fueled: Give diff and similarity between events	0.978	0.030	1.213	0.025	0.000	0.000
N130401	Fueled: Give suggestion about good value change	1.576	0.045	0.978	0.016	0.000	0.000
N130801	Cost to raise child: Find information from article	0.735	0.027	-1.012	0.043	0.000	0.000

Table 9-2d. Document literacy item descriptions and parameters for the National Adult Literacy Survey

Item#	Description	A	SE(A)	B	SE(B)	C	SE(C)
SCOR100	Social Security card: Sign name on line	0.504	0.025	-4.803	0.248	0.000	0.000
SCOR300	Driver's license: Locate expiration date	0.917	0.025	-2.525	0.058	0.000	0.000
AB20101	Energy graph: Find answer for given conditions (1)	1.153	0.045	-0.193	0.054	0.228	0.030
AB20201	Energy graph: Find answer for given conditions (2)	0.935	0.030	-0.023	0.045	0.096	0.023
AB20301	Energy: Yr 2000 source percent power larger than 71	1.089	0.036	0.684	0.031	0.142	0.015
AB20401	Yellow pages: Find a list of stores	0.478	0.019	-0.467	0.111	0.144	0.036
AB20501	Yellow pages: Find telephone number of given place	0.414	0.017	-0.771	0.111	0.088	0.031
AB20601	Yellow pages: Find place open Saturday	1.077	0.034	-0.143	0.041	0.105	0.023
AB20701	Bus schd: Take correct bus for given condition (1)	0.521	0.024	0.293	0.106	0.130	0.035
AB20801	Bus schd: Take correct bus for given condition (2)	1.282	0.044	0.901	0.024	0.144	0.012
AB20901	Bus schd: After 2:35, how long til Flint&Acad bus	1.168	0.032	1.520	0.021	0.162	0.008
AB21001	Bus schd: Take correct bus for given condition (4)	0.730	0.031	0.520	0.066	0.144	0.026
AB21501	With graph, predict sales for spring 1985	0.799	0.024	-0.571	0.038	0.000	0.000
AB30101	Street map: Locate intersection	0.953	0.027	-0.956	0.036	0.000	0.000
AB30301	Sign out sheet: Respond to call about resident	0.904	0.025	-0.844	0.034	0.000	0.000
AB30401	Sign out sheet: Respond to call about resident (2)	0.665	0.017	-0.089	0.028	0.000	0.000
AB30701	Major medical:locate Eligibility from table	0.960	0.026	-0.702	0.030	0.000	0.000
AB30801	Almanac: Find page containing chart for given info	0.704	0.017	0.929	0.019	0.000	0.000
AB30901	Almanac: Determine pattern in exports across years	0.299	0.013	0.000	0.056	0.000	0.000
AB31001	Abrasive guide: Type of sandpaper for sealing	0.831	0.019	0.285	0.020	0.000	0.000
AB31101	Abrasive gd: Can product be used in given case?	0.761	0.020	-0.256	0.028	0.000	0.000
AB31301	Facts about fire: Mark information in article	0.721	0.024	-1.170	0.055	0.000	0.000
AB40101	School registration: Mark correct age information	0.820	0.024	-1.063	0.041	0.000	0.000
AB40401	Almanac: Find page containing chart for given info	1.108	0.023	0.717	0.013	0.000	0.000
AB50401	Catalog order: Order product one	0.772	0.022	-0.882	0.039	0.000	0.000
AB50402	Catalog order: Order product two	0.771	0.018	0.396	0.019	0.000	0.000
AB50501	Telephone bill: Mark information on bill	0.359	0.014	-0.511	0.060	0.000	0.000
AB50601	Almanac football: Locate page of info in almanac	1.001	0.023	-0.083	0.020	0.000	0.000
AB50701	Almanac football: Explain why an award is given	1.182	0.029	-0.373	0.022	0.000	0.000
AB50801	Wage & tax statement: What is current net pay?	0.733	0.025	-1.365	0.060	0.000	0.000
AB50901	Wage & tax statement: What is yr-to-date gross pay	0.884	0.022	-0.199	0.025	0.000	0.000
AB60101	Make out check: Enter correct date on check	1.254	0.031	-0.497	0.021	0.000	0.000
AB60102	Make out check: Paid to the correct place	1.408	0.035	-0.425	0.018	0.000	0.000
AB60103	Make out check: Enter correct amount in numbers	0.993	0.026	-0.674	0.028	0.000	0.000
AB60104	Make out check: Enter correct amount written out	1.537	0.040	-0.524	0.018	0.000	0.000
AB60301	Phone message: Write correct name of caller	1.454	0.054	-1.283	0.036	0.000	0.000
AB60302	Phone message: Write correct number of caller	1.068	0.038	-1.434	0.048	0.000	0.000
AB60303	Phone message: Mark "please call" box	0.903	0.024	-0.680	0.030	0.000	0.000
AB60304	Phone message: Write out correct message	0.895	0.019	0.461	0.017	0.000	0.000
AB60305	Phone message: Write who took the message	0.640	0.017	-0.220	0.030	0.000	0.000
AB60306	Phone message: Write whom message is for	0.947	0.027	-0.867	0.033	0.000	0.000
AB60501	Petroleum graph: Label axes of graph	1.102	0.024	1.937	0.019	0.000	0.000
AB60502	Petroleum graph: Complete graph including axes	1.081	0.023	0.782	0.014	0.000	0.000
AB60701	Nurses' convention: Who would be asked questions	1.179	0.045	-1.295	0.047	0.000	0.000

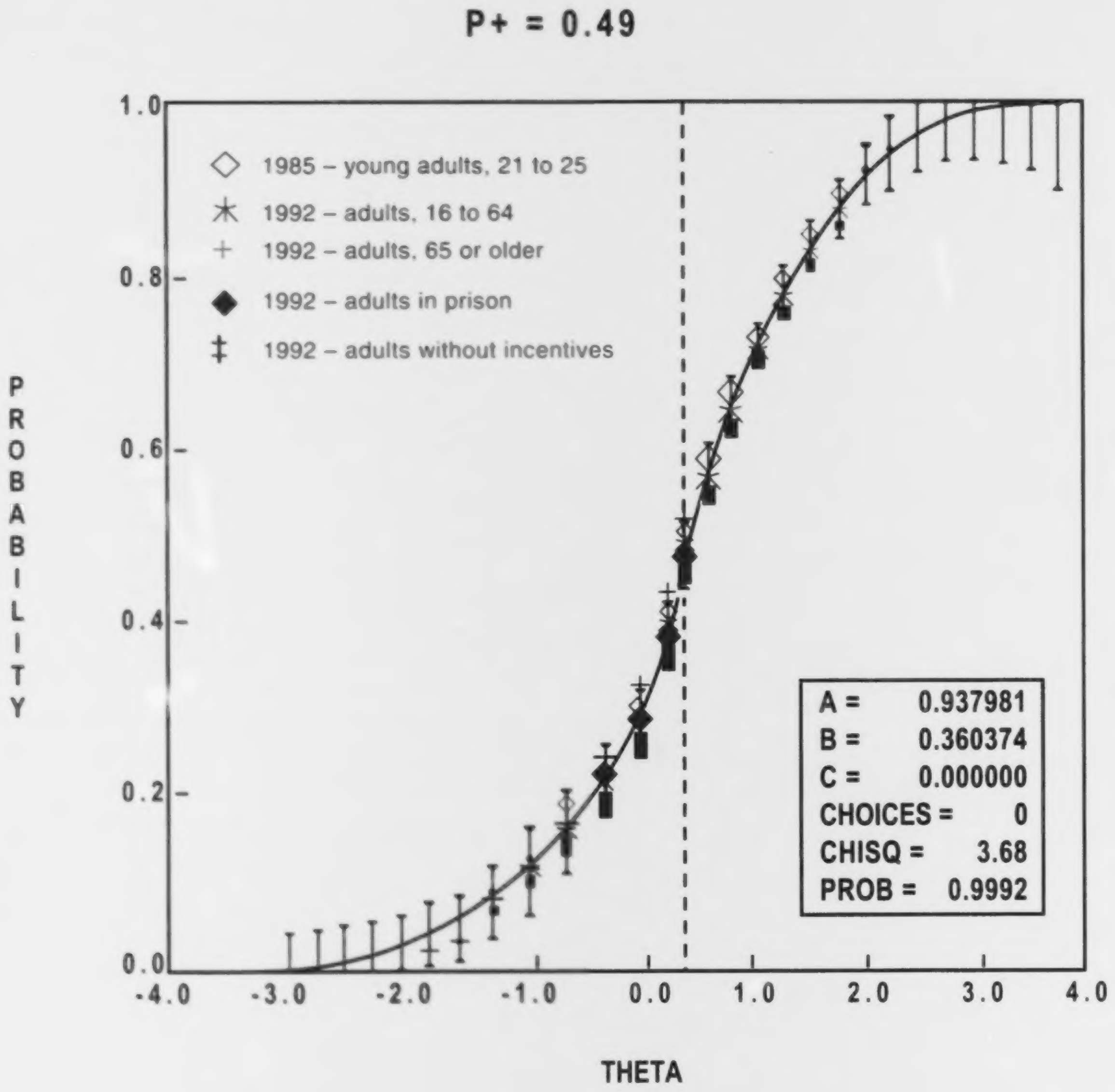
Table 9-2d. Document literacy item descriptions and parameters for the National Adult Literacy Survey — Continued

Item#	Description	A	SE(A)	B	SE(B)	C	SE(C)
AB60801	Nurses' convention: Write correct day of program	1.016	0.042	-1.539	0.063	0.000	0.000
AB60802	Nurses' convention: What is date of program?	1.231	0.058	-1.620	0.064	0.000	0.000
AB60803	Nurses' convention: What is time of program?	1.438	0.076	-1.649	0.063	0.000	0.000
AB61001	Nurses' convention: Write correct place for tables	0.766	0.030	-1.454	0.069	0.000	0.000
AB70104	Job application: Complete personal information	0.542	0.024	-2.337	0.119	0.000	0.000
AB70301	Almanac vitamins: Locate list of info in almanac	0.753	0.018	-0.134	0.025	0.000	0.000
AB70701	Follow directions on map: Give correct location	0.799	0.019	-0.126	0.024	0.000	0.000
AB70801	Classified: Match list with coupons	1.142	0.033	-0.880	0.030	0.000	0.000
AB70901	Checking deposit: Enter correct date	0.990	0.030	-1.088	0.039	0.000	0.000
AB70902	Checking deposit: Enter correct cash amount	0.858	0.021	-0.303	0.025	0.000	0.000
AB70903	Checking deposit: Enter correct amount of check	1.266	0.038	-0.921	0.029	0.000	0.000
AB71201	Mark correct movie from given information	0.939	0.041	-1.801	0.077	0.000	0.000
N010401	Vehicle chart: Find correct information	0.902	0.038	-1.340	0.062	0.000	0.000
N010801	Trend chart: Mark information on chart	0.807	0.028	-0.463	0.038	0.000	0.000
N010901	Trend chart: Put information on chart	0.720	0.024	1.702	0.032	0.000	0.000
N011001	Trend chart: Determine least # of points needed	0.645	0.022	0.260	0.032	0.000	0.000
N080601	Bus schedule: Take correct bus for given condition	1.039	0.029	0.505	0.020	0.000	0.000
N080701	Bus schedule: Mark map correctly for given info	1.094	0.034	-0.312	0.027	0.000	0.000
N080801	Auto maintenance form: Enter information given (1)	0.763	0.023	0.569	0.025	0.000	0.000
N080802	Auto maintenance form: Enter given information	1.357	0.048	-0.683	0.029	0.000	0.000
N090301	Essence: Determine page certain article begins on	1.123	0.048	-1.224	0.051	0.000	0.000
N090401	Essence: Determine topic of given article	0.987	0.033	-0.448	0.032	0.000	0.000
N090501	Essence: Determine topic of section of magazine	0.671	0.024	-0.301	0.040	0.000	0.000
N100501	Opinions table: Mark sentence explaining action	1.038	0.029	0.486	0.020	0.000	0.000
N100601	Opinions table: Find correct group for given info	1.134	0.032	1.284	0.019	0.000	0.000
N100701	Summarize views of parents & teachers	1.127	0.034	2.300	0.032	0.000	0.000
N110301	Certified mail rec't: Enter name and address	0.811	0.029	-0.742	0.045	0.000	0.000
N110302	Certified mail rec't: Enter postage and fee	0.714	0.028	-1.025	0.059	0.000	0.000
N110701	Credit card table: Find correct bank	0.469	0.020	0.125	0.047	0.000	0.000
N110901	Credit card table: Give 2 differences	0.829	0.031	1.882	0.032	0.000	0.000
N120101	Campus map: Mark map for given info	0.985	0.036	-0.801	0.040	0.000	0.000
N120201	Campus map: Find correct room for given dean	0.842	0.028	-0.403	0.035	0.000	0.000
N120601	Middle class: Find projected percent	0.795	0.037	-1.488	0.077	0.000	0.000
N130101	S.S. card application: Identify and enter info(1)	1.619	0.049	-0.095	0.017	0.000	0.000
N130102	S.S. card application: Identify and enter info(3)	1.270	0.043	-0.544	0.028	0.000	0.000
N130103	S.S. card application: Identify and enter info(2)	2.105	0.071	-0.290	0.016	0.000	0.000
N130104	S.S. card application: Identify and enter info(4)	2.159	0.069	-0.111	0.014	0.000	0.000

Table 9-2q. Quantitative literacy item descriptions and parameters for the National Adult Literacy Survey

Number	Description	A	SE(A)	B	SE(B)	C	SE(C)
AB40201	Unit price: Estimate cost/oz of peanut butter	0.818	0.019	0.455	0.017	0.000	0.000
AB40301	Unit price: Mark economical brand	0.815	0.034	0.216	0.029	0.447	0.000
AB40501	Airline schedule: plan travel arrangements (1)	0.909	0.020	0.005	0.016	0.000	0.000
AB40601	Airline schedule: plan travel arrangements (2)	0.952	0.021	-0.522	0.018	0.000	0.000
AB40701	Check ledger: Complete ledger (1)	1.597	0.034	-0.500	0.013	0.000	0.000
AB40702	Check ledger: Complete ledger (2)	1.936	0.042	-0.344	0.010	0.000	0.000
AB40703	Check ledger: Complete ledger (3)	1.873	0.040	-0.331	0.011	0.000	0.000
AB40704	Check ledger: Complete ledger (4)	1.970	0.042	-0.294	0.010	0.000	0.000
AB50301	Interest charges: Orally explain computation	0.601	0.020	1.522	0.043	0.000	0.000
AB50403	Catalog order: Order product three	0.609	0.016	0.600	0.023	0.000	0.000
AB50404	Catalog order: Shipping, handling, and total	0.968	0.023	-0.951	0.022	0.000	0.000
AB60901	Nurses Convention: Write number of seats needed	0.504	0.015	-0.355	0.031	0.000	0.000
AB70501	Lunch: Determine correct change using info in menu	0.893	0.019	0.090	0.016	0.000	0.000
AB70601	Lunch: Determine 10% tip using given info	0.872	0.019	0.384	0.016	0.000	0.000
AB70904	Checking deposit: Total bank deposit entry	0.869	0.029	-1.970	0.049	0.000	0.000
NC00501	Enter total amount of both checks being deposited	0.661	0.017	-2.792	0.060	0.000	0.000
NC00601	Price for Sleuth: how much less than On the Town	0.717	0.013	-1.690	0.028	0.000	0.000
N010501	Vehicle chart: Find sum of percentages	0.851	0.026	-0.768	0.029	0.000	0.000
N010601	Vehicle chart: Describe solution to percent problem	1.121	0.032	0.717	0.019	0.000	0.000
N010701	Vehicle chart: Find magnitude of difference	1.033	0.029	0.411	0.019	0.000	0.000
N011101	Gas gauge: Use info to answer question-show calcs	1.034	0.030	0.195	0.019	0.000	0.000
N080401	SSI: Calculate yrly amount for couple w/ basic ssi	0.696	0.022	0.520	0.026	0.000	0.000
N080501	Minutes from student union to 17th & Main	0.757	0.023	-0.247	0.025	0.000	0.000
N080901	Auto maintenance form: Calculate miles per gallon	0.850	0.027	0.856	0.026	0.000	0.000
N081001	Rank juices by expense and give reasons	0.732	0.023	0.122	0.025	0.000	0.000
N090101	Get discount if oil bill paid in 10 days	1.346	0.037	-0.018	0.016	0.000	0.000
N090201	Get net total owed after deduction	1.677	0.047	-0.349	0.015	0.000	0.000
N090901	Carpet ad: Get diff in reg and sale price	0.789	0.028	-1.003	0.040	0.000	0.000
N091001	Carpet ad: Get total cost to carpet room	0.634	0.026	1.371	0.045	0.000	0.000
N100801	Salt River: Determine difference in costs	0.647	0.027	-1.737	0.068	0.000	0.000
N100901	Salt River: Determine miles between stops	0.622	0.022	-0.263	0.032	0.000	0.000
N101001	Salt River: Determine hours between points	0.943	0.031	-0.837	0.031	0.000	0.000
N110201	Blood pressure: Calculate death rate from info	1.033	0.030	0.740	0.021	0.000	0.000
N110303	Certified mail rec't: Calculate postage and fees	0.789	0.031	-1.730	0.056	0.000	0.000
N110801	Credit card table: Determine difference in rates	0.881	0.029	-0.494	0.029	0.000	0.000
N120701	Calc percent diff black & white middle class-1980	0.909	0.029	-0.845	0.029	0.000	0.000
N120801	Middle class: Find difference in magnitude of pct	1.013	0.030	0.830	0.022	0.000	0.000
N121001	Calc miles/day Butcher went in this year's race	1.017	0.031	0.217	0.020	0.000	0.000
N121101	Susan Butcher: Calc diff in times for completion	0.959	0.035	1.517	0.035	0.000	0.000
N130501	Rec room: Calculate feet of molding needed	0.655	0.023	0.819	0.032	0.000	0.000
N130601	Rec room: Calculate number of wall panels needed	1.111	0.031	-0.184	0.019	0.000	0.000
N130701	Rec room: Describe solution of calculation needed	0.845	0.034	1.962	0.052	0.000	0.000
N130901	Raise child: Calc money needed to raise child	0.945	0.030	0.499	0.022	0.000	0.000

Exhibit 9-1. Item response curve for a task included in both the 1985 Young Adult Literacy Assessment and the 1992 National Adult Literacy Survey



9.3 PROFICIENCY ESTIMATION USING PLAUSIBLE VALUES

9.3.1 Generating Proficiency Scores

The purpose of most cognitive skills testing is to accurately assess individual performance for the purposes of diagnosis, selection, or placement. Regardless of which measurement model is being used, classical test theory or item response theory, the accuracy of these measurements can be improved—that is, the amount of measurement error can be reduced—by increasing the number of items given to the individual. Thus, achievement tests containing more than 70 items are common. Since the uncertainty associated with each θ is negligible, the distribution of θ or the joint distribution of θ with other variables can be approximated using individual θ's.

When analyzing the distribution of proficiencies for a group, however, more efficient estimates can be obtained from a sampling design like the one used in the National Adult Literacy Survey. The survey solicits relatively few responses from each sampled respondent while maintaining a wide range of content representation when responses are summed for all respondents. The advantage of estimating population characteristics more efficiently is offset by the inability to make precise statements about individuals. Uncertainty associated with individual θ estimates is too large to be ignored. Point estimates of proficiency that are, in some sense, optimal for each sampled respondent could lead to seriously biased estimates of population characteristics (Wingersky, Kaplan, and Beaton, 1987).

Plausible values methodology was developed as a way to estimate key population features consistently and to approximate others at the level of item response theory procedures. Mislevy (1991) provides a detailed review of plausible values methodology. Along with theoretical justifications, Mislevy presents comparisons with standard procedures, discusses biases that arise in some secondary analyses, and offers numerical examples.

The following is a brief overview of the plausible values approach, focusing on its implementation in the 1992 National Adult Literacy Survey analyses.

Let y represent the responses of all sampled respondents to background questions and questions on engagement to literacy activities, and let θ represent the scale proficiency values. If θ were known for all sampled examinees, it would be possible to compute a statistic $t(\theta, y)$—such as a scale or composite subpopulation sample mean, a sample percentile point, or a sample regression coefficient—to estimate a corresponding population quantity T.

Because the scaling models are latent variable models, however, θ values are not observed even for sampled respondents. To overcome this problem, we follow Rubin (1987) by considering θ as "missing data" and approximate $t(\theta, y)$ by its expectation given (x, y), the data that actually were observed, as follows:

$$\begin{aligned} t^*(\mathbf{x},\mathbf{y}) &= E[t(\theta,\mathbf{y})|\mathbf{x},\mathbf{y}] \\ &= \int t(\theta,\mathbf{y})p(\theta|\mathbf{x},\mathbf{y})d\theta \end{aligned} \tag{4}$$

It is possible to approximate t^* using random draws from the conditional distribution of the scale proficiencies given the item responses x_j, background variables y_j, and model parameters for sampled respondent j. These values are referred to as imputations in the sampling literature, and as plausible values in the National Adult Literacy Survey and in the National Assessment of Educational Progress. The value of θ for any respondent that would enter into the computation of t is thus replaced by a randomly selected value from his or her conditional distribution. Rubin (1987) proposed to repeat this process several times so that the uncertainty associated with imputation can be quantified by "multiple imputation." For example, the average of multiple estimates of t, each computed from a different set of plausible values, is a numerical approximation of t^* of the above equation; the variance among them reflects uncertainly due to not observing θ. It should be noted that this variance does not include the variability of sampling from the population.

It cannot be emphasized too strongly that plausible values are not test scores for individuals in the usual sense. Plausible values are only intermediary computations for calculating integrals as shown in the above equation in order to estimate population characteristics. When the underlying model is correctly specified, plausible values will provide consistent estimates of population characteristics, even though they are not generally unbiased estimates of the proficiencies of the individuals with whom they are associated. The key idea lies in a contrast between plausible values and the more familiar ability estimates of educational measurement that are in some sense optimal for each respondent (e.g., maximum likelihood estimates, which are consistent estimates of a respondent's θ, and Bayes estimates, which provide minimum mean-squared errors with respect to a reference population). Point estimates that are optimal for individual respondents have distributions that can produce decidedly nonoptimal (inconsistent) estimates of population characteristics (Little and Rubin, 1983). Plausible values, on the other hand, are constructed explicitly to provide consistent estimates of population effects. For further discussion, see Mislevy, Beaton, Kaplan, and Sheehan (1992).

Plausible values for each respondent j are drawn from the multivariate normal conditional distribution $P(\underline{\theta}_j|x_j,y_j,\Gamma,\Sigma)$, where Γ is a matrix of regression coefficients and Σ is a common variance matrix for residuals. Using standard rules of probability, the conditional probability of proficiency can be represented as follows

$$\begin{aligned} P(\theta_j|x_j,y_j,\Gamma,\Sigma) &\propto P(x_j|\theta_j,y_j,\Gamma,\Sigma)P(\theta_j|y_j,\Gamma,\Sigma) \\ &= P(x_j|\theta_j)P(\theta_j|y_j,\Gamma,\Sigma) \end{aligned} \tag{5}$$

where $\underline{\theta}_j$ is a vector of three scale values, $P(x_j \mid \theta_j)$ is the product over the scales of the independent likelihoods induced by responses to items within each scale, and $P(\underline{\theta}_j \mid y_j, \Gamma, \Sigma)$ is the multivariate joint density of proficiencies of the scales, conditional on the observed value y_j of background responses and parameters Γ and Σ. Item parameter estimates are fixed and regarded as population values in the computation described in this section. (See Appendix C for Γ (Gamma) values.)

In the National Adult Literacy Survey analyses, a normal multivariate distribution was assumed for $P(\theta_j \mid y_j, \Gamma, \Sigma)$, with a common variance, Σ, and with a mean given by a linear model with slope parameters, Γ, based on the first approximately principal components of several hundred selected main effects and two-way interactions of the complete vector of background variables. The background variables included sex, ethnicity, Spanish language interview, region of the country, respondent education, parental education, occupation, and reading practices. The complete set of original background variables used in the analyses is listed in Appendix G. Based on the principal component method, components representing 99 percent of the variance present in the data were selected. The included principal components will be referred to as the conditioning variables, and denoted as y^c. The following model was fit to the data:

$$\theta = \Gamma^c + \varepsilon \qquad (6)$$

where ε is normally distributed with mean zero and variance Σ. As in a regression analysis, Γ is a matrix each of whose columns is the effects for one scale and Σ is the three-by-three matrix variance of residuals between scales.

Note that in order to be strictly correct for all functions Γ of θ, it is necessary that $p(\theta \mid \mathbf{y})$ be correctly specified for all background variables in the survey. In the National Adult Literacy Survey, principal component scores were generated from background variables. Marginal means and percentile points of θ for these variables can be consistently estimated. Estimates of functions T involving background variables not conditioned in this manner are subject to error due to misspecification. The nature of these errors was discussed in detail in Mislevy (1991). Their magnitudes diminish as each respondent provides more cognitive data—that is, responds to a greater number of items. Indications are that the magnitude of these errors is negligible in the National Adult Literacy Survey (e.g., biases in regression coefficients below 5 percent) due to the larger numbers of cognitive tasks presented to each respondent in the survey (on average, 13 tasks per scale). The exception is the sample of respondents who could not or did not proceed beyond the background questions.

These respondents did not attempt the assessment tasks due to an inability to read or write English, a physical disability, a mental disability, or a refusal to participate in the survey. Chapter 8 describes the procedure used to estimate the proficiencies of those with missing responses. If these respondents had been excluded from the survey, the proficiency scores of some subpopulations in the National Adult Literacy Survey would have been severely overestimated, and the picture of the nation's literacy skills would have been distorted. These respondents possess few literacy skills, and detailed analyses of their proficiencies, not surprisingly, may lead to unstable results.

The basic method for estimating Γ and Σ with the EM procedure was described in Mislevy (1985) for a single scale case. The EM algorithm requires the computation of the mean, θ, and variance, Σ, of the posterior distribution. For the multiple scales of the National Adult Literacy Survey, the computer program C-GROUP (Thomas, 1993) was used. The program implemented a method to compute the moments using higher order asymptotic corrections to a normal approximation. Case weights were employed in this step.

After completing the EM algorithm, the plausible values are drawn in a three-step process from the joint distribution of the values of Σ for all sampled respondents with more than four cognitive tasks attempted. First, a value of Γ is drawn from a normal approximation to $P(\Gamma, \Sigma \mid x_j, y_j)$ that fixes Σ at the value $\hat{\Sigma}$ (Thomas, 1993). Second, conditional on the generated value of Γ (and the fixed value of $\Sigma = \hat{\Sigma}$), the mean θ, and variance Σ_j^P of the posterior distribution are computed using the same methods applied in the EM algorithm. In the third step, the θ values are drawn independently from a multivariate normal distribution with mean θ and variance Σ_j^P. These three steps are repeated five times, producing five imputations of θ for each sampled respondent.

For those with an insufficient number of responses, the Γ and Σs described in the previous paragraph were fixed. Hence, all respondents—regardless of the number of tasks attempted—were assigned a set of plausible values for the three scales. The plausible values can then be employed to evaluate an arbitrary function T according to the following five steps:

1. Using the first vector of plausible values for each respondent, evaluate T as if the plausible values were the true values of θ. Denote the result T_1.

2. In the same manner as in step 1 above, evaluate the sampling variance of T, or $\mathrm{Var}(T_1)$, with respect to respondents' first vectors of plausible values. Denote the result Var_1.

3. Carry out steps 1 and 2 for the second through fifth vectors of plausible values, thus obtaining T_u and Var_u for $u=2,\ldots,5$.

4. The best estimate of T obtainable from the plausible values is the average of the five values obtained from the different sets of plausible values:

$$T = \frac{\sum_u T_u}{5} \tag{7}$$

5. An estimate of the variance of T. is the sum of two components: an estimate of $\mathrm{Var}(T_u)$ obtained as in step 4 and the variance among the T_us:

$$\mathrm{Var}(T.) = \frac{\sum_u \mathrm{Var}_u}{5} + \left(1+\frac{1}{5}\right) \sum_u (T_u - T.)^2 \tag{8}$$

The first component in Var(T.) reflects uncertainty due to sampling respondents from the population; the second component reflects uncertainty due to the fact that the θs of the sampled respondents are not known precisely, but only indirectly through x and y.

9.3.2 Linking the 1992 Scale to the 1985 Scale

At this point, plausible values are still on the provisional scale and must be transformed to the 1985 scale for comparison. The 1985 scale was established in the following manner. In the 1985 assessment, some of the tasks administered were the same as those included in the NAEP 1984 reading assessment. Relying on the common tasks from the two assessments, the 1985 sample proficiency distribution was placed on the NAEP reading scale, a 0 to 500 metric. The mean and standard deviation of the plausible values for the 1985 samples were estimated to be 296.6 and 49.0, respectively. The mean and standard deviation of the other three scales—prose, document, and quantitative—were also set to these values.

In the 1992 National Adult Literacy Survey, as noted earlier, item parameters from the 1985 young adult literacy assessment were re-estimated using a larger sample and more accurate procedures than were available at the time of the 1985 analysis. These new item parameters are best suited for comparing performance distributions for different samples. However, the new sets of item parameters on the provisional scales and the old transformation constants used to produce the 1985 scales would not necessarily produce identical results for the 1985 sample. Thus, new linear transformation constants for the 1985 sample were found to match the mean and standard deviation of the current plausible value distribution of the 1985 sample based on the new item parameters. The same constants were applied to the 1992 sample proficiency distribution. The transformation that was applied is as follows: $\theta = A\theta^* + B$ where θ^* is the provisional scale from item calibration and θ is the reported 0 to 500 scale. Table 9-2 presents the transformation constants (that is, the standard deviations and means) for the distributions of the three scales. These constants apply both to the 1992 data, and to the 1985 data when the new item parameters are used.

Table 9-2. Transformation constants (standard deviations and means) by literacy scale, 1992 and 1985 (using new item parameters)

Literacy scale	A (standard deviations)	B (means)
Prose	51.67	269.16
Document	52.46	237.50
Quantitative	54.41	276.87

9.3.3 Evaluation of Differential Group Performance

Performance differences across subpopulations were examined by constructing empirical characteristic curves of tests rather than of items for major subpopulations defined by variables such as gender and ethnicity.

Yamamoto and Muraki (1991) have found that sets of estimated item parameters, each estimated on separate calibration samples with different racial/ethnic compositions, differed significantly even after an appropriate linear transformation was applied to account for the scale indeterminacy. This suggests differential item functioning (DIF) by racial/ethnic subpopulations. The National Adult Literacy Survey assessment as a whole functioned equivalently, however, suggesting that the effects of a different set of item parameters on the estimated proficiency of subpopulations may be negligible. In fact, after a linear scale transformation to account for the scale indeterminacy was applied to the real data, the estimates of subgroup proficiency distributions using a different set of item parameters were virtually identical. Since the main goal was to prevent systematic bias against any particular subpopulation, it was more appropriate to evaluate differential group performance at the test level than at the item level. Therefore, empirical test characteristic curves were constructed for the various sex, racial/ethnic, and age groups. These are shown in Exhibits 9-2p, 9-2d, and 9-2q, one for each scale.

The plots illustrate the average empirical proportion correct for the tasks in each literacy scale for each sex, racial/ethnic, and age group. Each point on the scale was estimated in two steps. First, the empirical proportion correct for every task was calculated for each sample for those whose proficiency values were in the selected 20-point range for at least one of 10 plausible values; second, the percents correct were then averaged for all tasks in the scale. This procedure was repeated for each subpopulation of interest. While the plot for document literacy scale by age groups (Exhibit 9-2d), and several others show deviations in the test characteristic curves within either the very low (below 200) and very high (above 360) parts of the proficiency ranges, the number of individuals performing in these ranges is very small, and therefore stable estimates cannot be made. Thus, when comparing test characteristic curves, one should concentrate on the part of the proficiency range where most of the population scores.

If the test characteristic curves deviated systematically within a subpopulation of interest, this could be viewed as evidence that the test is functioning differentially (is biased) for that group. The subpopulation curves were quite similar, however. Thus, it is safe to conclude that viewing the test as a whole, differential functioning was not observed across sex or racial/ethnic or age subpopulations in the National Adult Literacy Survey.

Exhibit 9-2p. Prose literacy test characteristic curves, by gender, race/ethnicity, and age: 1992

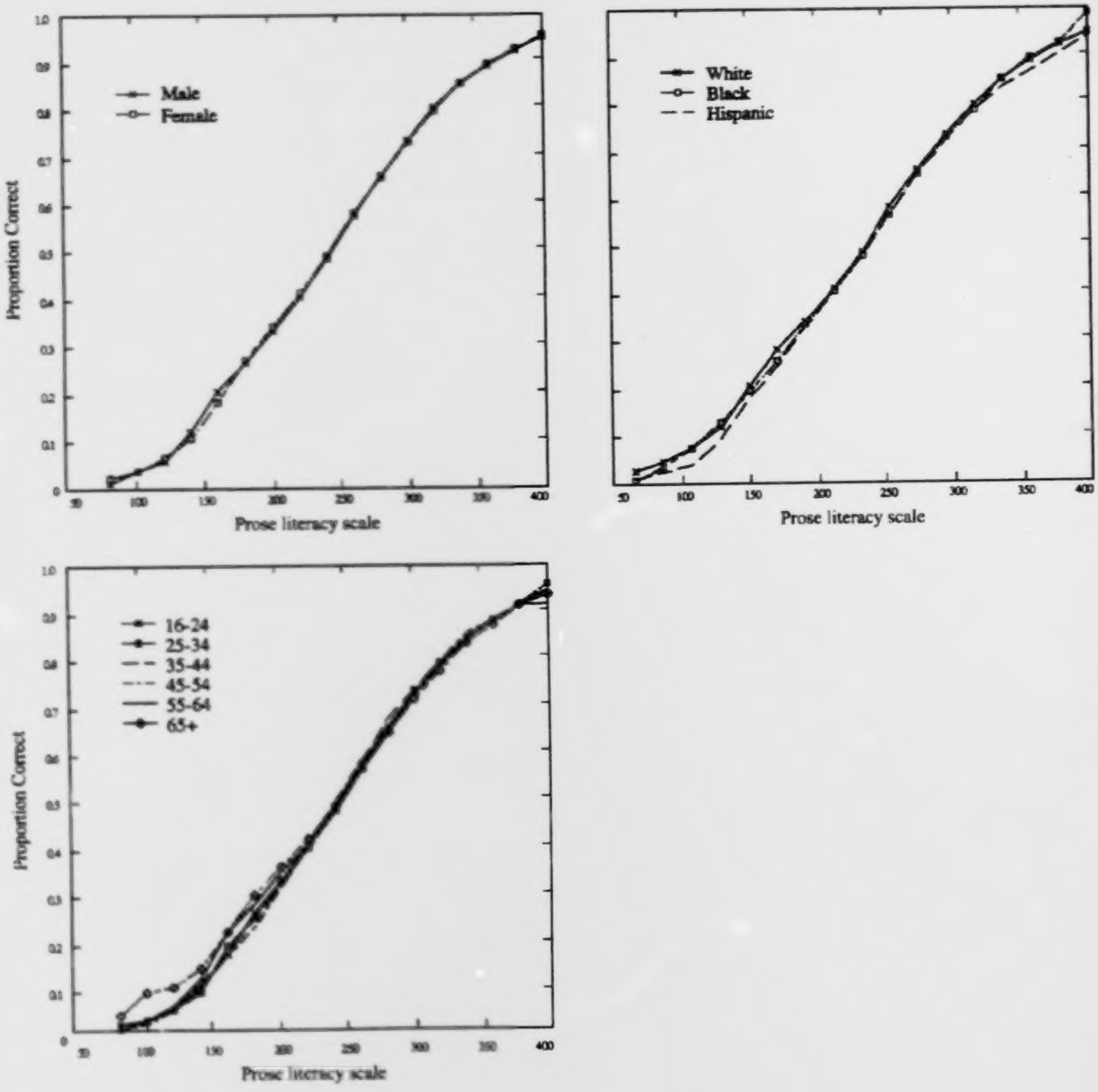


Exhibit 9-2d. Document literacy test characteristic curves, by gender, race/ethnicity, and age: 1992

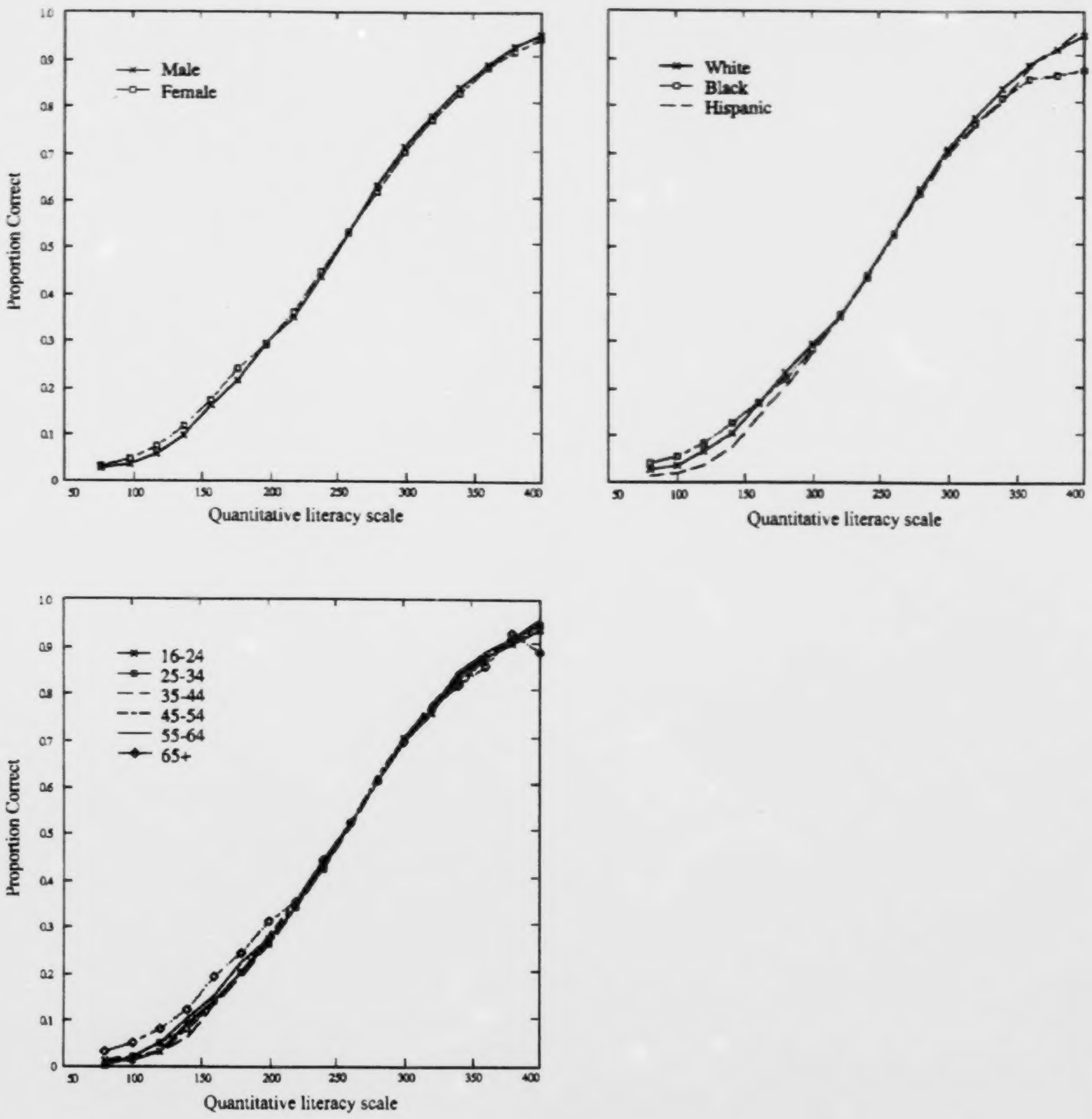
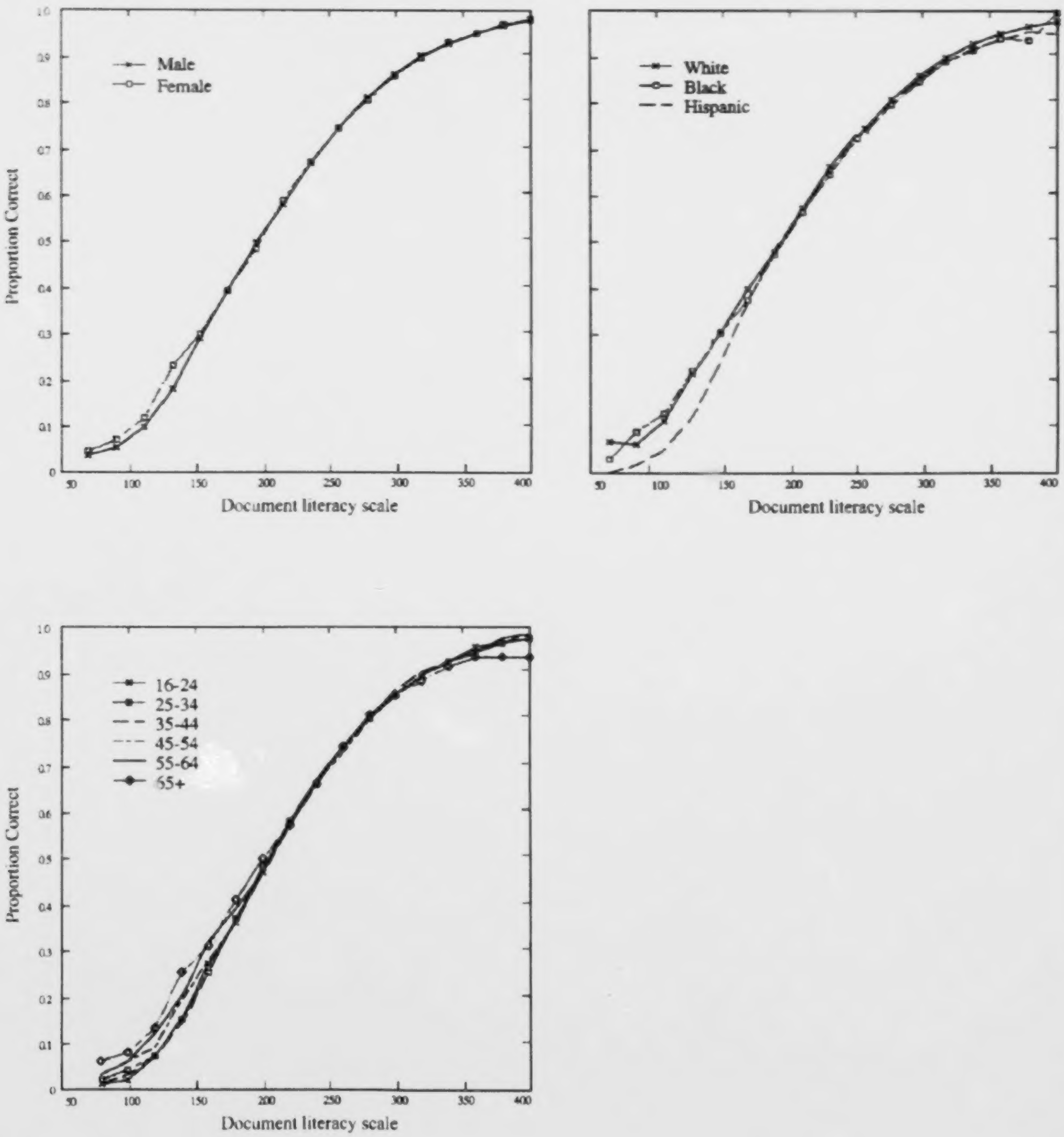


Exhibit 9-2q. Quantitative literacy test characteristic curves, by gender, race/ethnicity, and age: 1992



9.4 STATISTICAL TESTS

9.4.1 Analysis of Plausible Values

Plausible values methodology was used in this survey to increase the accuracy of the proficiency distribution estimates for various subpopulations and for the adult population as a whole. This method correctly retains the uncertainty associated with proficiency estimates for individual respondents by using multiple imputed proficiency values rather than assuming that this type of uncertainty is zero—a more common practice. Retaining this component of uncertainty requires that additional analysis procedures be used to estimate respondents' proficiencies.

If the true θ values were observed for all sampled respondents, the statistic $\frac{t - T}{\sqrt{U}}$ would follow a t-distribution with d degrees of freedom. Since the true θ values are unknown, only incomplete data are available. The corresponding incomplete-data statistic $\frac{t^* - T}{\sqrt{Var(t^*)}}$ is approximately t-distributed, with degrees of freedom given by

$$v = \frac{1}{\frac{f_M^2}{M - 1} + \frac{(1 - f_M)^2}{d}} \tag{9}$$

where f_M = the proportion of total variance due to not observing q values:

M = sets of plausible values

d = degrees of freedom associated with $\frac{t - T}{\sqrt{U}}$

$$f_M = \frac{\left(1 + \frac{1}{M}\right) B_M}{V_M} \tag{10}$$

where B_M = variance among the M estimates.

When B_M is small relative to U^* (average sampling variance over the M sets of plausible values), the reference distribution for incomplete-data statistics differs little from the reference distribution for the corresponding complete-data statistics. This was the case for the National Assessment of Educational Progress surveys. If, in addition, d is large, the normal approximation can be used instead of the t-distribution.

For k-dimensional t, such as the k coefficients in a multiple regression analysis, each U_M and U^* is a covariance matrix, and B_M is an average of squares and cross-products rather than simply an average of squares. In this case, the quantity $(T-t^*)V^{-1}(T-t^*)'$ is approximately F distributed with degrees of

$$f_M = \frac{(I + M^{-1})\text{Trace}(B_M V_M^{-1})}{k} \tag{11}$$

freedom equal to k and v, with v defined as above but with a matrix generalization of f_M

A chi-square distribution with k degrees of freedom can be used in place of f_M for the same reason that the normal distribution can approximate the t distribution.

Statistics t*, the estimates of ability and background variables, are consistent estimates of the corresponding population values T, as long as background variables are included in the conditioning variables. The consequences of violating this restriction are described by Beaton and Johnson (1990), Mislevy (1991), and Mislevy and Sheehan (1987). To avoid such biases, the National Adult Literacy Survey analysis included nearly all background variables, coded as dummy variables. To capture most of the variances in the background questions with a limited number of variables, principal components were used. Because each subpopulation can have unique relationships among the background variables, one set of principal components is not sufficient for all samples included in the National Adult Literacy Survey (i.e., the older adult, prison, and household samples). Each set of principal components was selected to include 99 percent of the variance in the background variables. Mislevy (1990) shows that this puts an upper bound of 1 percent on the average bias for all analyses involving the original conditioning variables.

9.4.2 Partitioning the Estimation Error Variance: A Numerical Example

This section offers an example of the use of multiple plausible values in the National Adult Literacy Survey analysis to partition the error variance. Table 9-3 presents data for three subgroups of respondents with differing educational attainments: those whose highest level of education was a GED, a high school diploma, and a four-year college degree. As noted earlier, five plausible values were calculated for each respondent for each scale. Each column presents the means of these five values.

Table 9-3. Mean plausible values by level of education for the prose scale

Level of Education	Sample N	Five imputed values					Mean	Var	JK_1 var	Standard error
		1	2	3	4	5				
GED	1062	269.3	268.1	267.9	268.2	267.7	268.2	0.483	2.888	1.84
High school	6107	270.2	270.4	270.3	270.5	270.2	270.3	0.180	1.050	1.11
4-year college	2534	321.2	321.7	322.4	322.8	320.4	321.7	1.027	1.408	1.56

Variance in the mean plausible values is similar but not identical for the three subgroups. As noted previously, variance reflects a component of error attributable to the measurement instrument's lack of precision and a component of error attributable to sample size. Variance can be reduced by either increasing the precision of the measurement instrument (for example, expanding the number of items) or increasing the size of the sample. The jackknife method was used to estimate error variance due to sampling using the first set of imputed values. This component of variance is expected to be consistent across the imputed values, and the size is influenced by the homogeneity of proficiencies among respondents in a subgroup but not by the sample size or by the precision of the survey instruments. Error variance due to sampling is smaller when the subgroup consists of respondents with similar proficiencies.

Despite a relatively large sample size, the mean for respondents with four-year college degrees has a larger error variance than those for other education groups. In fact, it is twice as large as the variance for respondents whose highest level of education is a GED. The higher variance for this best educated group is due to the characteristics of the assessment, which encompassed the entire adult population (age 16 and older) in this country and measured a wide range of skills. The precision of the assessment is optimal at the middle of the proficiency range, since that is where most of the population is expected to perform. Since the majority of the respondents with four-year college degrees scored above this range, variance due to lack of precision in measurement is quite high. Therefore, increasing the sample size would not do much to reduce the variance component for this group. On the other hand, the error variance due to sampling is twice as large for the smaller GED group as for the larger four-year college degree group.

The last column presents the standard error of the subpopulation mean, which is equal to the square root of the sum of the two components of error variance. The differences among the means can be compared using these standard errors. In doing so, it is first necessary to decide how many comparisons are being made. For this example, one might be interested in making three comparisons: GED vs. high school, high school vs. four-year college degree, and GED vs. four-year college degree. Following the Bonferroni method of multiple comparisons, any comparison among these three with a standardized difference greater than 2.39—$(\text{mean}_1 - \text{mean}_2)/\text{sqrt}(se_1^2 + se_2^2)$, ($z_p = 0.025/3$)—can be considered statistically significant. The difference in means between GED recipients and high-school graduates is not statistically significant at the .05 level, but the differences between these two groups and respondents with four-year degrees are significant.

9.4.3 Minimum Sample Sizes for Reporting Subgroup Results

In the National Adult Literacy Survey reports, the sample sizes were not always large enough to permit accurate estimates of proficiency and/or background results for one or more categories of variables. For results to be reported for any subgroup, a minimum sample size of 45 was required. This number was arrived at by determining the sample size needed to detect an effect size of 0.5 with a probability of 0.8 or greater using a design effect of 1.5. This design effect implies a sample design-based variance 1.5 times that of simple random sampling. The effect size of 0.5 pertains to the true difference in mean proficiency between the subgroup in question and the total population, divided by the standard deviation of proficiency in the total population. An effect size of 0.5 was chosen following Cohen (1988), who classifies effect size of this magnitude as "medium."

9.4.4 Estimates of Standard Errors with Large Mean Squared Errors

Standard errors of mean proficiencies, percentages, and percentiles play an important role in interpreting subpopulation results and comparing the performances of two or more subpopulations. The jackknife standard errors reported for the National Adult Literacy Survey are statistics whose quality depends on certain features of the samples from which the estimates are obtained. In certain cases—primarily when the standard error is based on a small number of respondents—the mean squared error associated with the estimated standard errors may be quite large. In the survey reports, estimated standard errors that are subject to large mean squared errors are followed by the symbol "!", indicating that the coefficient of variation (CV) is greater than 0.2. This CV is estimated by:

$$\mathrm{CV}(\hat{N}) = \frac{\mathrm{SE}(\hat{N})}{\hat{N}} \qquad (12)$$

where $\hat{N}$ is a point estimate of N and SE($\hat{N}$) is the jackknife standard error of $\hat{N}$.

Experience with other large-scale assessments suggests that when this coefficient exceeds 0.2, the mean squared error of the estimated standard errors of means, and percentages based on samples of this size, may be quite large. Therefore, these standard errors, and any confidence intervals or significance tests involving them, should be interpreted with caution. Johnson and Rust (1992) discuss this issue in detail.

Chapter 10

THE ROLE OF INCENTIVES IN LITERACY SURVEY RESEARCH

Leyla Mohadjer, Martha Berlin, Susan Rieger, Joseph Waksberg, Westat, Inc.;
Donald Rock, Kentaro Yamamoto, Irwin Kirsch, ETS; Andrew Kolstad, NCES

Before conducting the National Adult Literacy Survey, an experimental study was carried out under the survey contract to test the effect of monetary incentives on response rates, test performance, and survey costs. This study was part of a more general field test of data collection methods planned for the National Adult Literacy Survey. The experiment indicated that incentives increased response rates substantially, particularly among minorities and persons with low educational levels. Furthermore, moderate incentives ($20 per participant) actually lowered survey costs, mostly because of the reduction in the number of return visits to the household to obtain participation. On the basis of these findings, the $20 incentive was paid in the survey. However, there was some uncertainty about whether the findings of an experiment, carried out by a small number of interviewers and supervisors, would apply to a national survey involving over 400 interviewers and a more decentralized field structure. Because of the importance of the issue for both future cycles of the National Adult Literacy Survey and similar surveys, it was decided to replicate the main features of the experiment in the survey. The experimental features of the field test and the national study and the results of the research follow. Also presented is a brief summary of the literature on the use of monetary incentives in survey research.

10.1 LITERATURE REVIEW

The widespread use of survey research today as a means of gathering information makes its quality a matter of considerable interest. Because response rates are important features of quality, concerned researchers have explored a variety of response inducement techniques over the past 25 years. In one extensive effort to determine which methods of response rate improvement are most effective, Kanuk and Berenson (1975) examined over 75 articles that addressed techniques for increasing mail survey response rates. Given the large number of studies reviewed, they found that follow-up contact and the use of monetary incentives were the only two methodological procedures that had any empirical impact on response rates. Groves and Couper (1998) provide a theoretical framework for understanding and studying household survey nonresponse rates. They explore the influences that effect nonresponse rates, including ways to reduce nonresponse error in survey estimates.

Before introducing monetary incentives as a response rate improvement technique in the National Adult Literacy Survey, a literature review was conducted. Most reported monetary incentive experiments have focused on mail questionnaire surveys and, to a lesser extent, on telephone surveys. Few have focused on surveys conducted in person. Furthermore, many have been limited to special populations such as physicians (Berry and Kanouse, 1987; Gunn and Rhodes, 1981), other professionals (Godwin, 1979), college students (Zusman and Duby, 1987), and prison inmates (Novak et al., 1977).

Research has found that the use of monetary incentives does increase response rates in mail surveys. In support of this finding, Armstrong (1975) reviewed 18 studies by 14 researchers and concluded not only that monetary incentives in mail surveys yield large increases in response rates, but that larger incentives produce greater increases in response. In 27 cases examined by Armstrong (1975) and Linsky (1975), response rates were increased by an average of 20 percent through the use of monetary incentives, and in six studies the response rate for the group receiving the incentive was more than twice that of the control group.

In a review of literature on improving survey response rates, Baxter et al. (1984) discussed the theory that most respondents need a reason for taking the time to participate in a survey. In other words, the rewards for an action must outweigh the costs if the action is to take place. For most surveys, the rewards for participation are intangible ones, such as positive regard for the respondent or an opportunity to participate in a socially useful survey. The study by Baxter et al. indicates that for every respondent there is an amount of money that will function as a symbolic reward and engender participation.

Several reports on the effect of monetary incentives on response rates in personal interview studies are consistent with the research on mail surveys. For example, the Centers for Disease Control and the National Opinion Research Center conducted an experiment to determine the effectiveness of paying monetary incentives to physicians for participating in a 20- to 30-minute interview. The interview was to be conducted either in person or by telephone, depending on the doctor's preference. Six hundred physicians were systematically assigned to one of three subsamples designated to receive no incentive, \$25, or \$50. The response rate for the group that was offered no incentive was approximately 58 percent. Response rates increased to 69 percent for the \$25 incentive group and to 77 percent for the \$50 incentive group (Gunn and Rhodes, 1981).

The only study reviewed that did not show improvement in response rates with the use of incentives is a 1972-73 household survey conducted by the U.S. Bureau of the Census for the Bureau of Labor Statistics (Walsh, 1977). In the Consumer Expenditure Survey, an experiment was undertaken to determine the effects on response rates of offering a modest cash incentive to sample households for maintaining a record of their expenses for a 2-week period. The overall sample was split into three subsamples, one group receiving no monetary incentive, one group receiving \$5, and the third group receiving \$10. (Using the Consumer Price Index, with the 1993 dollar as standard, \$5 and \$10 in 1972 are equivalent to approximately \$15 and \$30, respectively, in 1993.) Overall, response rates ranged from 72.4 percent for the \$0 group to 76.9 percent for the \$10 group, a difference of only 4.5 percentage points. This finding led the researchers to conclude that incentives did not appear to be the most cost-effective technique for improving cooperation. However, design and operational problems flawed the experiment, and the Census Bureau noted that the results might have been different if the experiment had been conducted after these problems had been addressed.

Despite the outcome of the Census Bureau experiment, most of the evidence on the use of monetary incentives has pointed favorably toward this form of response inducement (Duncan, 1979). Most researchers have found that monetary incentives can prevent unsuccessful survey outcomes resulting from poor response rates and can improve sample representativeness (Zusman and Duby, 1987).

The precursor to the 1992 National Adult Literacy Survey, the 1985 Young Adult Literacy Assessment, was conducted by Educational Testing Service and Response Analysis Corporation as part of the National Assessment of Educational Progress (NAEP). Literacy assessments were completed with 3,600 respondents in their homes, during a six-month period in 1985. A response rate of 80.5 percent was achieved. Each assessment included a 30-minute background and attitude questionnaire, followed by a 60-minute battery of tasks designed to measure literacy skills. A $15 check was given to each respondent who completed the questionnaire and the assessment.

The Workplace Literacy Survey, a Department of Labor study based on the young adult literacy assessment framework, was administered in 1990 to eligible adults enrolling in the Job Training Partnership Act program, served by the Employment Service, or applying for benefits through the Unemployment Insurance program. This survey also included the administration of a background questionnaire and literacy tasks. To ensure comparability with the young adult literacy assessment results, each respondent received, upon completion of the survey, a payment comparable to the amount paid in the 1985 Young Adult Literacy Assessment.

10.2 THE 1991 FIELD TEST

10.2.1 Field Test Design

The 1991 field test for the National Adult Literacy Survey was designed to study the effectiveness of a variety of procedures planned for the survey. A principal purpose was to examine the effect of monetary incentives on response rates, test performance, and survey costs. The field test was conducted in a sample of 16 primary sampling units (PSUs), consisting of counties or groups of counties representing the 48 contiguous states. The PSUs were selected in a way that satisfied a Latin Square design based on key variables thought to be related to response rate. The variables included region of the country, urbanicity, race/ethnicity, and the average income/education level of persons residing in the selected PSUs. On average, 21 segments (consisting of census blocks) were selected within each PSU, with about eight households selected in each segment.

Incentive levels of $0, $20, and $35 were to be evaluated in the 1991 field test. The payment of $20 was proposed because it was comparable to the amount paid in the NAEP 1985 young adult literacy assessment, adjusted for inflation. An upper limit of $35 was selected for comparison with the $20 payment. Incentives were randomly assigned to segments so that each incentive group had about the same number of cases, and approximately the same number of incentive groups were represented at the PSU and census region levels.

Thirty-eight interviewers carried out the survey, in most cases two per PSU. The interviewers were assigned approximately equal numbers of segments in the three payment groups. The survey included three data collection instruments: a screener, a background questionnaire, and an exercise booklet. The screener collected household-enumeration data for the purpose of selecting an eligible respondent. The background questionnaire collected information in six areas: demographic data, language background, education, political and social participation, labor force participation, and literacy activities. The exercise booklet consisted of three 15-minute sections of prose, document, and quantitative tasks. Interviewers were instructed to introduce the incentive after the household composition had been determined and the eligible respondent(s) selected.

Respondents who completed the background questionnaire and agreed to complete the exercise booklet were given the incentive check after completing the exercise booklet. In order to compare the costs associated with the three incentive levels (as well as response rates and other measures of data quality), interviewers were required to record time spent and expenses incurred by segment each week.

The field test design included the administration of about 100 new literacy tasks (or nine blocks assembled into nine exercise booklets), and these tasks had to be evenly distributed across the three incentive levels. The goal was to distribute the booklets as indicated in Table 10-1.

To accomplish this distribution, the booklets were preassigned to households. This plan permitted the statistical evaluation of both the impact of incentives and the effect of block position.

10.2.2 Summary of Field Test Results

The analyses of the field test data showed that incentive payments significantly improved the outcomes of the survey. (The term significantly in this document is used in the statistical sense that is, any difference would not have resulted from chance alone.)

Table 10-1. Configuration and distribution of exercise booklets for the field test, by incentive group

		Number of booklets			
Booklet number	Blocks	$0 incentive	$20 incentive	$35 incentive	Total
1	A B C	74	74	74	222
2	B C A	74	74	74	222
3	C A B	74	74	74	222
		Total number of booklets prepared per block or task			666
4	D E F	74	74	74	222
5	E F D	74	74	74	222
6	F D E	74	74	74	222
		Total number of booklets prepared per block or task			666
7	G H I	74	74	74	222
8	H I G	74	74	74	222
9	I G H	74	74	74	222
		Total number of booklets prepared per block or task			666
Total number of booklets prepared per incentive group		666	666	666	1998

The following major areas showed significant improvement as a result of incentive payments:

- *Response Rates.* Incentives produced significant increases in response rates for the background questionnaire and exercise booklets (incentives were not introduced with the screener). Statistically significant differences in response rates were not detected between the $20 and $35 incentive groups.

- *Representation of the Target Population.* Analyses showed that the incentives were most effective in improving response rates for people with low educational attainment and for minority populations who are frequently underrepresented in national household surveys (the issues related to undercoverage in household surveys are discussed later). As a result, the use of incentives resulted in a better representation of the general population and of subgroups, such as the black and Hispanic populations, that were of special interest to the National Adult Literacy Survey.

- *Relationship Between Incentive Levels, Self-Selection, and Performance.* The accuracy of the literacy estimates depends on gaining the cooperation of the majority of eligible respondents. A significant number of refusals within any one of the three incentive levels would have biased the results unless the respondents who refused did not differ in any relevant or systematic way from the respondents who completed the exercises in the remaining incentive levels. Results of the field test showed that if an incentive payment had not been offered to eligible respondents, the main assessment would have had substantially increased bias in estimates of the population's literacy level, because self-selection factors occurring in the $0 incentive group would have had a deleterious effect on the representativeness of the sample. If incentives had not been used, self-selection factors resulting from reduced response rates among persons with lower levels of educational attainment would have overestimated the literacy levels in the United States for both the total population and major subgroups. The poststratification procedures used in the survey partially compensate for such factors, but poststratification would probably not have completely eliminated the apparent bias.

- *Survey Costs.* A cost analysis of the field test experiment showed a reduction in interviewing costs when incentives were given to the respondents. The cost per completed interview (including the cost of the incentive) for respondents in the \$20 incentive group was lower than in the \$0 and \$35 incentive groups.

A more in-depth discussion of the results of the 1991 field test with respect to the impact of incentives follows.

10.2.3 Field Test Response Rates

Of the 2,774 households in the sample, 336 units (12.1 percent) were either vacant or did not satisfy the definitions of a dwelling unit at the time of screening. As a result, 2,438 households were found to be eligible to participate in the survey, of which 2,155 (88.4 percent) completed the screener. Of 283 nonresponding households, 152 (6.2 percent) refused to participate in the study, and 131 (5.4 percent) did not complete the screener for reasons such as language problems and health reasons, etc.

The screener response rate was computed as follows:

$$\text{Screener response rate} = \frac{\text{Number of completed screeners}}{\text{Number of sampled dwelling units Vacant/Not a dwelling unit}} \qquad (1)$$

In households with completed screeners, one adult was randomly selected if there were three or fewer eligible adults in the household. Two adults were selected in households with four or more eligible people. After the interviewer selected the respondent(s), a background questionnaire was administered to each respondent. A Spanish translation of the background questionnaire was provided to bilingual interviewers to administer to Spanish-speaking respondents. However, the literacy exercises were written only in English.

The background questionnaire (BQ) response rate was computed as follows:

$$\text{BQ response rate} = \frac{\text{Number of completed BQs}}{\text{Number of persons selected at screener}} \qquad (2)$$

Upon completion of the background questionnaire, the exercise booklet was administered. The exercise booklet (EX) response rate was computed in the following way:

$$\text{EX response rate} = \frac{\text{Number of completed EXs}}{\text{Number of persons completing BQs — Ineligibles}^{*}} \qquad (3)$$

*Ineligibles mean cases with a language barrier, physical disability, mental disability, or respondent could not read.

Table 10-2 presents the number of respondents who were eligible to complete the background questionnaire and the exercise booklet. From the 2,155 households completing the screener, 2,288 eligible respondents were selected for the background questionnaire and the exercise booklet. Of these, 412 did not complete the background questionnaire; 258 (62.6 percent) of these were refusals, 39 (9.5 percent) had some type of mental or physical disability, and the remainder were nonrespondents for other reasons (such as unavailability during the field period or the interviewer's inability to contact the respondent despite repeated attempts).

The largest group of nonrespondents were persons who refused to complete the background questionnaire.

The response rate analyses indicated a statistically significant improvement in the background questionnaire response rates for respondents given incentives. The response rate increased by approximately 4 percentage points when an incentive of $20 was paid. There were no statistically significant improvements in response rate when the incentive was increased from $20 to $35. Appendix R presents the standard errors and t-values associated with the analyses performed in this chapter.

Table 10-2. Number of respondents completing the background questionnaire and exercise booklet in the field test, by incentive group

	$0 incentive	$20 incentive	$35 incentive	Overall
Background questionnaire (BQ)				
Number of persons sampled	730	740	818	2,288
Complete	574	612	690	1,876
Refusal	90	83	85	258
Other nonresponse	66	45	43	154
Response rate (percent)	78.6	82.7	84.4	82
Exercise booklet				
Number of BQ respondents	574	612	690	1,876
Complete	490	557	634	1,681
Ineligible*	46	43	43	132
Partial complete*	2	0	4	6
Refusal	36	12	9	57
Response rate (percent)	92.8	97.9	98	96.4

*Respondents in the ineligible category included those who were physically or mentally incapable of responding, those who could not read, and those who were not proficient in English. The partial complete category includes respondents who started the exercise but did not complete it.

All 1,876 persons who completed the background questionnaire were asked to complete an exercise booklet. Among the different incentive levels, the same pattern of response rates was observed for the exercise booklet as for the background questionnaire. The analysis showed a significant improvement in response rates for respondents given incentives. The 5 percentage points increase in exercise booklet response rates when respondents were paid a $20 or $35 incentive is statistically significant. There are, however, no statistically significant differences in response rates between the $20 and $35 incentive groups.

10.2.4 Representation of the Target Population in the Field Test

When a monetary incentive was paid, a disproportionate share of the increase in response rates occurred in population subgroups that are of special interest to the National Adult Literacy Survey, that is, among minorities and persons without a college education. This effect of incentive payments is at least as important as the improvement in overall response rates. The sample design involved oversampling black and Hispanic adults to allow for more detailed analyses of these subgroups, which together account for more than 20 percent of the current U.S. population.

It is particularly important to achieve high response rates for minorities to partially compensate for the undercount that almost always occurs for these populations in household surveys. Evaluations of the decennial censuses indicate that there has been a recurrent undercount of approximately 2 percent to 3 percent of the population (Citro and Cohen, 1985). Furthermore, since at least 1950, coverage of black persons in censuses has been considerably lower than coverage of White persons, with the proportion of uncounted black males higher than for other major demographic subgroups. Recent studies of undercounts have shown that the coverage of Hispanic persons is probably even lower than the coverage of black persons. Sample surveys usually achieve even lower rates of coverage than censuses. Unpublished data for the principal household surveys conducted by the Bureau of the Census indicate much lower coverage rates than in the census and that the coverage has been decreasing in the last 10 to 20 years. Consequently, low response rates would intensify even further the potential coverage bias.

The analyses of the field test data indicate that incentives significantly increased the response rates, particularly in the black population. The influence of incentives was similar for the Hispanic and non-Hispanic White populations with the Hispanic adults having the highest response rate, with or without incentives. In the non-incentive group there were quite large differences in response rates among the three race/ethnicity groups. These differences were sharply reduced in both the $20 and $35 incentive panels. Table 10-3 shows the response rates for the background questionnaire and exercise booklet for subgroups of interest. The table shows statistically significant ($p<.05$) improvements in response rates when the incentive was given to black respondents. There was a statistically significant increase of about 9 percentage points in the background questionnaire response rate for the minority population (black and Hispanic populations combined) as a result of providing a $20 incentive. The exercise booklet response rates for the minority populations also demonstrated a statistically significant increase of approximately 10 percentage points when a $20 incentive was offered. Thus, the overall effect of the $20 incentive was to add approximately 20 percentage points to the response rate for minorities. There were no statistically significant differences in response rates between the $20 and $35 incentive groups on the exercise booklet.

Persons with lower levels of educational attainment also represent an important subgroup from a public policy perspective. Table 10-3 includes the distribution of exercise booklet response rates by respondent education and incentive level. (The educational attainments of the sampled persons who did not complete the background questionnaire were not known; thus, the background questionnaire response rate could not be calculated by education of the sampled persons.) The table shows a significant increase in exercise booklet response rates in the $20 and $35 incentive groups, compared with the $0 group, for persons with only a high school diploma.

Table 10-3. Response rates for the background questionnaire and the exercise booklet in the National Adult Literacy Survey field test, by incentive level and age, race/ethnicity, and education

	Background questionnaire response rate					Exercise booklet response rate				
	N	$0	$20	$35	Overall	N	$0	$20	$35	Overall
Total selected	2,288	78.6	82.7	84.4	82.0	1,876	92.8	97.9	98.0	96.4
Age										
16-24	331	80.0	88.4	90.2	86.7	287	100.0	100.0	100.0	100.0
25-44	959	78.4	87.3	86.0	84.2	807	94.4	98.3	99.6	97.6
45-64	534	74.9	79.0	83.3	79.2	423	89.6	95.8	96.2	94.0
65-74	253	83.7	75.8	83.1	81.4	206	93.4	97.8	95.3	95.2
75+	160	86.0	73.6	84.2	81.3	130	74.3	97.0	91.9	87.6
Missing	51	54.5	35.7	46.2	45.1	23	100.0	33.3	100.0	87.5
Race/ethnicity										
White, non-Hispanic	1,769	78.7	81.1	85.3	81.9	1,448	93.4	97.5	97.9	96.4
Black, non-Hispanic	226	75.4	88.3	83.3	82.7	187	86.4	98.5	97.1	94.9
Hispanic	161	85.7	91.4	89.4	88.8	143	91.7	100.0	100.0	96.9
Other	69	80.0	95.8	84.0	87.0	62	90.9	100.0	100.0	98.0
Missing	63	66.7	57.1	57.1	60.3	36	91.7	66.7	100.0	90.0
Education										
Still in HS						68	100.0	95.5	100.0	98.5
No HS diploma						358	72.1	78.0	78.9	76.7
HS diploma						486	86.7	94.9	96.6	89.4
Some college						553	92.3	95.2	97.3	95.1
College degree						345	95.2	96.2	98.2	96.5
Missing						66	46.7	100.0	65.0	64.5

The same pattern exists for those with only a high school diploma. The exercise booklet response rate increased 6 percentage points for persons without a high school diploma, and 8 percentage points for persons with only a high school diploma when a $20 incentive was given. The increase in response rates for persons with some college education or a college degree was not significant. Again, no significant improvement in response rates was found as a result of increasing the incentives from $20 to $35.

The experiment also showed a significant increase in response rates for persons ages 16-64 when incentives were given: When respondents were paid $20, there was an increase of approximately 8 percentage points in background questionnaire response rates and an increase of approximately 4 percentage points in exercise booklet response rates. For the population age 65 and older, the effect of paying an incentive is not clear. The incentive did significantly increase the exercise response rate, but the impact on the background questionnaire was not significant.

10.2.5 Relationship Between Incentive Level, Self-Selection, and Performance in the Field Test

The previous discussions have focused on the differential effects of three levels of incentive payment on various response rates in the National Adult Literacy Survey field test. The analyses reported in this section were undertaken to investigate the impact of the three incentive conditions on the distributions of estimated proficiency scores on the three literacy scales—prose, document, and quantitative—as well as to identify position effects, if any, for the blocks of newly developed literacy tasks that were pretested.

An area of particular interest is whether the groups who agreed to complete the exercise booklets represented individuals of similar demonstrated literacy proficiencies; that is, is there evidence that individuals of higher (or lower) literacy proficiency were more likely to participate under a given incentive condition?

Before discussing the results of the analyses, it may be helpful to describe briefly the exercise booklets and the literacy score estimates derived from the field test. Some 100 literacy tasks were developed and assembled into nine discrete blocks, each of which was expected to require 15 minutes of administration time. Each block contained approximately the same number of tasks relevant to each of the three literacy scales. The nine blocks of tasks were then assembled into nine booklets, each requiring a total of approximately 45 minutes of administration time. The booklets were configured so that the same three blocks appeared together in three booklets, with each block placed in each position that is, as the first, middle, and last block of a booklet (see Table 10-1). The item response theory scaling procedure used to estimate literacy proficiency scores allows one to put all scores on a scale, even though groups of individuals complete different sets of tasks (Lord, 1980). The scales defined for the field test (prose, document, and quantitative) ranged from 0 to 100, with a mean of 50 and a standard deviation of 10.

The first concern is to examine the impact of incentive level on the response rate for the exercise booklet. Table 10-4 summarizes information about the relationship between incentive levels and gaining initial cooperation to administer the exercises. The eligible respondents referred to in Table 10-4 include all sampled individuals who completed a background questionnaire in English. The respondents who are classified as incomplete are those who did not respond to any of the literacy tasks; this classification includes respondents with physical disabilities, mental disabilities, or inadequate English language skills. (For these reasons, the figures in this table do not correspond exactly to the figures in Table 10-2.)

While response rates were generally high across all incentive levels, there were proportionately more incompletes at the $0 incentive level (12.6 percent) than at either of the other two incentive levels (6.6 percent and 5.1 percent, respectively). The differences between the response rates of the $0 incentive group and the other two groups were statistically significant. There was no statistical or practical difference in response rates between the $20 and $35 incentive levels.

Table 10-4. Initial cooperation rate, by incentive level: 1991 field test

	$0 incentive		$20 incentive		$35 incentive	
Total eligible	566	100.0%	603	100.0%	684	100.0%
Complete	495	87.4%	563	93.4%	649	94.9%
Incomplete	71	12.6%	40	6.6%	35	5.1%

Another consideration is the impact of the incentive levels on the cooperation of the elderly subsample. As shown in Table 10-5, eligible respondents age 65 and older who were assigned to the $0 condition were less likely to agree to complete the exercise booklet than their counterparts participating under the $20 (t = -3.1) and $35 (t = -3.0) incentive conditions.

Table 10-5. Initial cooperation rate, by incentive level: older adults in 1991 field test

	$0 incentive		$20 incentive		$35 incentive	
Total eligible	135	100%	88	100%	116	100%
Complete	103	76%	80	91%	104	90%
Incomplete	32	24%	8	9%	12	10%

One key question is whether or not the individuals who did cooperate under the $0 incentive level have different literacy proficiencies than those who cooperated under the other two incentive levels. To address this question, a two-way analysis of variance was conducted, where the independent variables are exercise booklets (three groups of booklets) and the three levels of incentives. The dependent variables are provisional literacy scale scores (prose, document, and quantitative). As outlined in Table 10-1, there were nine booklets, divided into groups of three. Each group of three booklets had a different set of three blocks of tasks, which were arranged so that each block appeared in each of the three possible positions. While the present design controlled for the effects of block position, it is interesting to note that a secondary analysis showed no statistically significant position effects. That is, respondents' demonstrated proficiencies did not vary according to the order in which the blocks appeared in their exercise booklets. As shown in Table 10-6, only incentive level was related to performance on the literacy scales. For all three dependent variables (that is, proficiencies on the three literacy scales), incentive level had a significant impact on performance.

The next important question concerns the nature of this impact of incentive level on literacy proficiency scores. Tables 10-7, 10-8, and 10-9 present literacy proficiencies on the prose, document, and quantitative literacy scales, respectively, by incentive level and by selected demographic characteristics crossed with incentive levels. Results from the analysis of variance show that incentive levels can account for differences in the estimated proficiency.

For all three literacy scales, the total mean scores for the $0 incentive level were significantly higher than the corresponding means for the $20 and $35 incentive levels. There was no significant difference between the total mean literacy scores of the $20 and $35 incentive levels on any of the scales. Further inspection of Tables 10-7 through 10-9 indicates that the performance difference in favor of the $0 incentive recipients is approximately 20 percent of a standard deviation on each of the three scales. In the education evaluation literature, a difference in group mean performance of greater than one-tenth of a standard deviation, "a small but nontrivial difference," is judged typically to be of practical educational significance (Cohen, 1988). Differences of the magnitude reported here, therefore, could lead to a serious positive bias in estimating the literacy proficiency for the adult population if no incentive were given in the main survey.

Table 10-6. Two-way analysis of variance on scale scores, by exercise booklet by incentive and interaction: 1991 field test data

Source	Sum of squares of deviations	Degrees of freedom	Mean square error	F statistic	P value
Prose literacy total	4,463,243.5	1,707			
Mean	4,231,233.6	1	4,231,233.6	41,702.9	0.0000
Books	32.8	2	16.4	0.2	0.8507
Incentives	1,340.4	2	670.2	6.6	0.0014
B*I	269.9	4	67.5	0.7	0.6162
Error	172,281.5	1,698	101.5		
Document literacy total	4,470,963.9	1,707			
Mean	4,243,240.2	1	4,243,240.2	42,577.6	0.0000
Books	10.7	2	5.4	0.1	0.9497
Incentives	1,338.1	2	669.0	6.7	0.0012
B*I	258.2	4	64.5	0.7	0.6286
Error	169,221.0	1,698	99.7		
Quantitative literacy total	4,441,101.6	1,707			
Mean	4,218,290.5	1	4,218,290.5	43,415.4	0.0000
Books	93.8	2	46.9	0.5	0.6173
Incentives	1,562.5	2	781.2	8.0	0.0003
B*I	147.2	4	36.8	0.4	0.824
Error	164,979.7	1,698	97.2		

Table 10-7. Prose proficiency sample sizes, means, and standard deviations, by incentive group: 1991 field test

	Total			$0 incentive			$20 incentive			$35 incentive		
	N	Mean	StdDv	N	Mean	StdDv	N	Mean	StdDv	N	Mean	StdDv
Total	1,707	50.1	(10.1)	495	51.4	(10.1)	563	49.3	(9.9)	649	49.8	(10.1)
Sex												
Male	775	50.2	(10.2)	225	51.8	(9.7)	271	48.8	(10.3)	279	50.2	(10.3)
Female	917	50.1	(10.0)	264	51.3	(10.4)	290	49.8	(9.5)	363	49.4	(9.9)
Race/ethnicity												
White	1,370	51.1	(10.0)	405	52.5	(10.0)	441	50.2	(10.1)	524	50.9	(9.9)
Black	171	44.8	(9.2)	39	44.9	(9.7)	60	44.9	(7.5)	72	44.5	(10.1)
Hispanic	124	46.7	(9.5)	43	49.0	(9.3)	45	46.9	(9.4)	36	43.7	(8.9)
Education												
Still in H.S.	68	47.3	(10.3)	20	47.2	(12.7)	21	47.6	(10.2)	27	47.2	(8.2)
< HS	94	42.1	(10.0)	20	43.5	(8.3)	30	38.1	(9.5)	44	44.2	(10.3)
Some HS	180	43.6	(9.3)	47	44.3	(9.2)	73	43.5	(8.9)	60	43.3	(9.9)
GED/HSEQ	54	46.0	(9.3)	11	45.0	(6.2)	22	44.0	(8.9)	21	48.6	(10.5)
HS diploma	405	48.2	(9.0)	102	49.8	(10.0)	147	47.3	(8.3)	156	48.0	(8.7)
Some college	438	51.9	(9.0)	137	51.8	(9.0)	131	52.2	(8.7)	170	51.7	(9.2)
College degree	437	55.8	(8.6)	143	57.2	(7.9)	133	55.6	(8.1)	161	54.7	(9.3)
Age												
16-20	156	48.6	(9.8)	44	48.7	(10.9)	48	47.1	(8.6)	64	49.5	(9.8)
21-25	153	51.8	(9.8)	42	51.8	(8.8)	45	52.4	(11.6)	66	51.4	(9.0)
26-31	211	53.0	(9.0)	48	54.0	(8.4)	82	51.8	(9.7)	81	53.9	(8.2)
32-45	544	51.5	(9.8)	155	53.3	(10.5)	202	50.2	(9.4)	187	51.3	(9.5)
46-64	356	49.5	(9.8)	103	51.0	(9.5)	106	49.2	(9.1)	147	48.7	(10.3)
65+	287	46.1	(10.6)	103	49.1	(10.2)	80	44.0	(9.8)	104	44.7	(10.8)
Income												
<$5,000	61	46.7	(8.5)	16	47.6	(8.5)	22	45.9	(6.8)	23	46.7	(9.7)
$5,000-9,999	98	45.4	(11.1)	20	42.8	(9.0)	28	47.9	(10.2)	50	45.0	(11.9)
$10,000-14,999	142	47.2	(9.9)	33	50.2	(10.3)	46	46.3	(10.1)	63	46.4	(9.3)
$15,000-19,999	128	48.8	(10.2)	33	49.8	(7.2)	47	46.3	(11.5)	48	50.6	(10.2)
$20,000-29,999	241	49.3	(10.2)	80	50.5	(10.5)	83	49.1	(10.2)	78	48.4	(9.8)
$30,000-39,999	237	51.2	(9.8)	52	53.6	(10.3)	80	50.2	(9.1)	105	50.7	(9.9)
$40,000-49,999	204	50.6	(8.6)	66	52.4	(8.5)	69	49.1	(8.8)	69	50.4	(8.1)
$50,000+	465	53.9	(9.1)	144	55.4	(8.9)	150	52.7	(9.5)	171	53.7	(8.8)
Refused	40	44.6	(10.8)	17	43.4	(12.1)	14	47.7	(9.1)	9	41.8	(9.1)
Don't know	64	45.1	(10.5)	22	44.9	(10.0)	17	43.0	(8.3)	25	46.8	(12.0)

Table 10-8. Document proficiency sample sizes, means, and standard deviations, by incentive group: 1991 field test

	Total			$0 incentive			$20 incentive			$35 incentive		
	N	Mean	StdDv	N	Mean	StdDv	N	Mean	StdDv	N	Mean	StdDv
Total	1707	50.2	(10.1)	495	51.6	(9.8)	563	49.8	(9.9)	649	49.5	(10.1)
Sex												
Male	775	50.6	(10.3)	225	51.9	(9.6)	271	49.9	(10.4)	279	50.1	(10.5)
Female	917	49.9	(9.8)	264	51.4	(10.0)	290	49.6	(9.5)	363	49.0	(9.8)
Race/ethnicity												
White	1370	51.2	(9.8)	405	52.2	(9.7)	441	50.7	(10.0)	524	50.8	(9.7)
Black	171	44.8	(9.1)	39	46.1	(8.5)	60	46.2	(8.4)	72	43.0	(9.6)
Hispanic	124	46.6	(10.8)	43	50.8	(9.7)	45	45.3	(9.7)	36	43.2	(11.8)
Education												
Still in H.S.	68	50.3	(10.5)	20	51.7	(11.1)	21	51.0	(9.2)	27	48.7	(10.6)
Less than HS	94	40.9	(11.3)	20	44.9	(11.1)	30	40.1	(10.9)	44	39.7	(11.3)
Some HS	180	44.3	(8.6)	47	45.0	(8.6)	73	44.3	(8.2)	60	43.9	(9.1)
GED/HSEQ	54	46.2	(8.5)	11	47.4	(4.8)	22	46.1	(7.4)	21	45.7	(10.8)
HS diploma	405	48.6	(8.6)	102	49.5	(8.2)	147	48.2	(9.0)	156	48.5	(8.3)
Some college	438	52.0	(9.0)	137	52.9	(9.2)	131	51.6	(9.3)	170	51.7	(8.6)
College degree	437	55.1	(8.9)	143	55.7	(8.9)	133	55.5	(8.6)	161	54.3	(9.1)
Age												
16-20	156	51.7	(9.7)	44	53.2	(9.8)	48	50.2	(8.9)	64	51.8	(9.9)
21-25	153	53.2	(8.3)	42	53.5	(7.5)	45	52.9	(7.2)	66	53.4	(9.4)
26-31	211	52.5	(9.3)	48	54.1	(8.9)	82	51.5	(9.7)	81	52.5	(8.9)
32-45	544	52.1	(9.0)	155	54.0	(9.1)	202	51.9	(9.1)	187	50.8	(8.8)
46-64	356	48.8	(10.2)	103	49.7	(10.3)	106	48.4	(10.0)	147	48.4	(10.4)
65+	287	44.1	(10.1)	103	47.2	(9.5)	80	42.5	(10.3)	104	42.4	(9.8)
Income												
<$5,000	61	45.6	(10.4)	16	49.8	(10.8)	22	45.4	(8.1)	23	42.9	(11.1)
$5,000-9,999	98	45.3	(10.1)	20	47.7	(8.3)	28	46.5	(9.2)	50	43.7	(11.0)
$10,000-14,999	142	47.2	(9.6)	33	50.7	(9.8)	46	45.4	(8.0)	63	46.7	(10.1)
$15,000-19,999	128	48.2	(9.3)	33	47.1	(7.1)	47	47.5	(10.1)	48	49.7	(9.5)
$20,000-29,999	241	48.5	(10.2)	80	49.8	(10.3)	83	48.4	(10.7)	78	47.4	(9.5)
$30,000-39,999	237	50.7	(9.4)	52	53.0	(9.0)	80	51.2	(8.6)	105	49.1	(9.9)
$40,000-49,999	204	51.2	(9.1)	66	52.6	(9.0)	69	49.6	(9.4)	69	51.4	(8.5)
$50,000+	465	54.3	(8.9)	144	55.0	(9.1)	150	54.0	(8.8)	171	54.1	(8.9)
Refused	40	46.3	(11.4)	17	45.7	(11.7)	14	48.3	(11.8)	9	44.3	(9.8)
Don't know	64	47.7	(10.2)	22	48.6	(8.9)	17	45.6	(12.6)	25	48.3	(9.3)

Table 10-9. Quantitative proficiency sample sizes, means, and standard deviations, by incentive group: 1991 field test

	Total			$0 incentive			$20 incentive			$35 incentive		
	N	Mean	StdDv	N	Mean	StdDv	N	Mean	StdDv	N	Mean	StdDv
Total	1707	50.0	(9.9)	495	51.5	(9.3)	563	49.5	(10.5)	649	49.4	(9.7)
Sex												
Male	775	51.4	(10.1)	225	52.8	(9.8)	271	50.7	(10.5)	279	50.9	(9.9)
Female	917	48.9	(9.5)	264	50.5	(8.7)	290	48.3	(10.3)	363	48.1	(9.2)
Race/ethnicity												
White	1370	51.2	(9.5)	405	52.6	(8.8)	441	51.1	(9.9)	524	50.4	(9.6)
Black	171	43.2	(8.7)	39	44.1	(8.6)	60	42.1	(9.1)	72	43.6	(8.2)
Hispanic	124	45.9	(9.7)	43	49.2	(9.7)	45	43.4	(10.1)	36	45.0	(7.9)
Education												
Still in H.S.	68	47.0	(9.8)	20	47.0	(9.4)	21	45.3	(10.8)	27	48.2	(9.1)
Less than HS	94	41.9	(10.2)	20	43.3	(7.5)	30	42.9	(10.8)	44	40.6	(10.6)
Some HS	180	43.2	(8.4)	47	44.8	(7.3)	73	41.5	(8.5)	60	44.4	(8.9)
GED/HSEQ	54	46.5	(6.7)	11	45.8	(5)	22	45.5	(7.5)	21	48.0	(6.3)
HS diploma	405	48.3	(9)	102	50.0	(8.8)	147	47.7	(9.6)	156	47.7	(8.5)
Some college	438	51.5	(8.7)	137	52.3	(8.4)	131	51.9	(9.1)	170	50.6	(8.5)
College degree	437	55.9	(8.6)	143	56.7	(8.2)	133	56.3	(8.9)	161	54.9	(8.6)
Age												
16-20	156	48.7	(9.3)	44	49.3	(8.9)	48	47.1	(10.8)	64	49.6	(8.3)
21-25	153	50.5	(9.8)	42	52.4	(9.6)	45	48.6	(11.8)	66	50.6	(7.9)
26-31	211	52.3	(9.2)	48	54.6	(8.2)	82	50.7	(9.4)	81	52.7	(9.2)
32-45	544	51.4	(9.6)	155	52.6	(9)	202	51.4	(10.3)	187	50.0	(9.1)
46-64	356	50.1	(9.8)	103	51.6	(9)	106	49.5	(10.5)	147	49.4	(9.7)
65+	287	46.1	(10.2)	103	49.0	(9.6)	80	45.4	(9.6)	104	43.8	(10.5)
Income												
<$5,000	61	44.6	(11.2)	16	46.6	(12.3)	22	43.5	(10.9)	23	44.4	(10.6)
$5,000-9,999	98	45.6	(9.2)	20	45.9	(8)	28	45.3	(8.3)	50	45.6	(10)
$10,000-14,999	142	46.8	(10.4)	33	48.5	(9.3)	46	44.6	(10.3)	63	47.4	(10.6)
$15,000-19,999	128	47.6	(10.9)	33	48.8	(9.2)	47	49.1	(12.7)	48	45.2	(9.6)
$20,000-29,999	241	48.7	(9.3)	80	49.6	(9)	83	48.7	(9.6)	78	47.7	(9.3)
$30,000-39,999	237	50.6	(9.2)	52	53.2	(9.7)	80	49.3	(8.6)	105	50.2	(9.1)
$40,000-49,999	204	52.1	(8)	66	52.6	(7.5)	69	52.1	(8.6)	69	51.8	(7.7)
$50,000+	465	54.1	(8.9)	144	55.7	(7.7)	150	53.8	(10.1)	171	53.1	(8.5)
Refused	40	48.5	(11.1)	17	49.7	(10.9)	14	48.2	(8.8)	9	46.7	(13.9)
Don't know	64	44.9	(8.7)	22	45.9	(7.3)	17	38.8	(8.7)	25	48.0	(7.6)

The apparent self-selection of higher scoring individuals in the $0 incentive group is further demonstrated by the fact that, compared with the $20 and $35 incentive groups, this group had a disproportionately greater percentage of individuals with some college education or a college degree. That is, when the education level data in Table 10-7, 10-8, or 10-9 are combined, 57 percent of all respondents in the $0 incentive group have some college education or a college degree, while the corresponding figures for the $20 and $35 incentive levels were 47 percent and 51 percent, respectively. Similarly, respondents in the $0 incentive group were proportionately more likely to come from households with relatively high income levels. About 42 percent of the $0 incentive respondents were in households reporting incomes of $40,000 or more, while the corresponding figures for the $20 and $35 incentive levels were 39 percent and 37 percent, respectively.

The issue of performance and self-selection is important to consider with respect to the 65-and-older population as well. There were 287 individuals age 65 or older who completed an exercise booklet in the field test. Those in the $0 incentive group scored significantly higher than their counterparts in the $20 and $35 incentive groups (see Tables 10-7 through 10-9). This result is consistent across all three literacy scales. While there were too few 65-and-older individuals in each of the incentive groups to permit an investigation of possible bias patterns related to differential self-selection, the significantly higher performance of the elderly $0 recipients suggests a positive selection bias similar to that found in other $0 incentive groups. As indicated above, this pattern of higher proficiency scores in the $0 incentive group is a consistent finding not only for the 65-and-older population but for virtually all other subpopulations on all three literacy scales.

Given the mean score differences by incentive level for the entire field test sample, it is apparent that if incentives had not been offered to eligible the National Adult Literacy Survey respondents, the sample for the main assessment would likely be unrepresentative, and the results would have overestimated the level of literacy in the United States. Moreover, as seen in Tables 10-7 through 10-9, the overestimation is likely to be an even greater problem in the results for a number of subgroups of interest. For example, the $0 incentive condition can be expected to yield skewed proficiency distributions for Hispanic adults; for individuals with 0-8 years of education and for high school and college graduates; for those over the age of 26 and, in particular, for senior citizens; and for individuals at all household or personal income levels, except possibly the $5,000 to $9,000 level. Although poststratification procedures might be used to partially compensate for this apparent bias, it is unlikely that it would have been completely eliminated.

Another question that is addressed using the field test data is whether or to what extent incentives affect respondents' motivation to perform. Table 10-10 presents the results of a two-way analysis of variance of booklet groups by incentive level, with proportion of items attempted as the dependent variable. To the extent that the number of items attempted is a proxy for motivation, this analysis attempts to evaluate the direct impact of incentive level on the respondent's motivation to perform after having made the decision to complete the exercise booklet. As shown in Table 10-10, there was no significant relationship between incentive level and proportion of items attempted. It appears from the data that response bias came from self-selection factors that affected whether or not an individual chose to participate in the assessment. Once an individual had agreed to participate, level of incentive did not seem to have an impact on the individuals motivation to perform as measured by the number of tasks attempted.

Table 10-10. Proportion of items attempted, by exercise booklet group and incentive level: Two-way analyses of variance, 1991 field test

Source	Sum of squares of deviations	Degrees of freedom	Mean square error	F statistic	P value
Total	1497.766	1,707			
Mean	1392.694	1	1392.694	29296.59	0.0000
Books	0.181	2	0.091	1.91	0.1487
Incentives	0.139	2	0.070	1.46	0.2317
B*I	0.114	4	0.029	0.60	0.6629
Error	80.719	1,698	0.048		

10.2.6 Survey Costs for the Field Test

One final but important component in evaluating monetary incentives is to review their impact on survey costs. Therefore, record-keeping procedures were implemented in the field test to allow the analysis of interviewer time and expense data by level of incentive payment.

Table 10-11 shows the level of effort and cost per completed assessment for each of the three incentive groups. The field test experiment indicated that the cost of conducting the assessment was reduced when a $20 incentive was paid to the respondent. This was true even when the cost of the incentive was added to the interviewer costs. When the $20 incentive was added to the interviewer costs, the net cost to the survey was $92.24, compared with $95.89 when no incentive was paid and $103.91 when the $35 incentive was paid. The cost per completed assessment includes only interviewer wages and expenses (mileage, telephone, tolls, etc.) and has not been adjusted to include overhead, general and administrative costs, and fee. The costs within an incentive group were divided by the number of completed assessments within the incentive group to calculate the average cost per completed assessment.

Table 10-11. Survey costs and level of effort by incentive group: 1991 field test

	Incentive group		
	$0	$20	$35
Average interviewer hours per completed assessment	8.4	6.8	6.4
Average number of contacts per completed assessment	6.7	5.3	5.0
Average interviewer costs per completed assessment	$95.89	$72.24	$68.91
Average interviewer costs + cost of incentive per completed assessment	$95.89	$92.24	$103.91

The primary reason for this result was that the incentive reduced the number of contacts (callbacks) the interviewer had to make to complete an assessment. The average number of contacts per completed assessment decreased from 6.7 when no incentive was paid to 5.3 or 5.0 contacts per completed assessment when a $20 or a $35 incentive was paid, respectively. The number of contacts required to gain respondent cooperation strongly influenced the hours per completed assessment, because interviewer travel time to and from respondents' homes and time spent in multiple attempts to convert nonrespondents are, in addition to questionnaire administration time, the key components of the hours to complete an assessment.

10.2.7 Conclusions from the 1991 Field Test

The analyses of the field test incentive experiment indicate that a $20 incentive significantly increased the response rate, especially for subgroups of particular interest to the study. Increasing the response rates for groups that are less likely to participate in the survey improved the distribution of the sample and the representation of the target population. Furthermore, the likelihood of producing a biased estimate of the population's literacy level increases when no incentive is offered. This outcome is related to the greater tendency for self-selection factors to occur in the non-incentive group, resulting in a nonrepresentative sample. More specifically, when there is no monetary incentive, better educated individuals and individuals with higher income levels (i.e., those tending to have higher levels of literacy proficiency) are more likely to agree to complete the literacy tasks, while those with lower levels of educational attainment are less likely to do so. As noted earlier, poststratification procedures could reduce the effect of this problem, but it is unlikely that they would completely eliminate the resulting biases. Once an individual has decided to complete the literacy exercises, however, the incentive level seems to have little or no effect on the person's motivation to respond, as measured by the number of tasks attempted.

The field test results also indicate that the cost of conducting the assessment is lower for the $20 incentive group than for the $0 and $35 incentive groups. The net cost per completed interview in the $20 condition is $92.24; for the $0 incentive condition it is $95.89. The primary reason for the lower cost is that the incentive reduced the number of return visits the interviewer had to make to complete an assessment.

In summary, the use of a $20 incentive in the 1991 National Adult Literacy Survey field test resulted in (1) a better response rate and, therefore, a larger number of completed assessments; (2) completed assessments from respondents who closely resemble the general population, thereby greatly reducing the likelihood of biased estimates of the population's literacy level; and (3) improvements in the survey with a lower cost to the government. Therefore, it was decided to pay a $20 incentive for the main 1992 data collection of the National Adult Literacy Survey.

The results of the field test were consistent with those of previous studies carried out to evaluate the use of incentives in surveys, as presented in the literature review (Section 10.1).

10.3 THE 1992 INCENTIVE EXPERIMENT

An incentive experiment similar to the one carried out in the field test was conducted as part of the national component of the 1992 National Adult Literacy Survey. The major purposes of the experiment were to ascertain the extent to which the field test results predicted the effect of an incentive in the large scale national survey and to improve the precision of the field test results by increasing sample sizes, particularly for the $0 incentive group. The analysis in this section is restricted to the effect of incentives on response rates. In order to improve the precision of the data, the incentive and non-incentive subsamples were interpenetrated to the extent practical, and the same interviewers handled both types of cases. Separate cost data for the two treatments were not maintained. However, the field test finding that a moderate incentive reduced field costs is likely to apply to the National Adult Literacy Survey because the incentive and non-incentive response rate differences were equivalent in the 1991 field test and the 1992 National Adult Literacy Survey.

10.3.1 Sample Design for the 1992 National Adult Literacy Survey

The target population for the national component of the National Adult Literacy Survey[1] was adults age 16 or older in the 50 states and the District of Columbia who, at the time of the survey (February through August, 1992), resided in private households.

A stratified sample design with the following stages was used:

Stage 1:	PSUs consisting of counties or groups of counties
Stage 2:	Secondary sampling units (SSUs), consisting of census blocks or groups of neighboring blocks (segments)
Stage 3:	Housing units
Stage 4:	Age-eligible individuals

The first-stage sample was a sample of PSUs, which were geographic clusters of one or more adjacent counties. In developing the sampling frame, the 3,141 counties and independent cities in the 50 states were grouped into 1,404 PSUs, from which a sample of 101 PSUs was selected.

The SSUs were segments that consisted of individual blocks or groups of neighboring blocks. Segment selection followed procedures commonly used for area sampling. Segments with high proportions of black or Hispanic residents were oversampled to increase the representation of these minority groups in the sample. The sample consisted of 2,064 segments.

---

[1] In addition to the national component of the National Adult Literacy Survey, about 1,000 adults were surveyed in each of eleven states that chose to participate in a special study designed to produce state-level results that are comparable to the national data. The State Adult Literacy Surveys were conducted concurrently with the National Adult Literacy Survey and were carried out by the same data collection staff.

The third stage of sampling involved the selection of 24,522 housing units from the listings developed by the field listers within the selected segments. (Listing quality control procedures conducted during the data collection effort resulted in the addition to the sample of 305 housing units, making a total of 24,827 households.)

In the fourth stage, 17,863 eligible persons within responding households were selected to complete the background questionnaire and exercise booklet.

10.3.2 1992 Incentive Experiment Design

A subsample of 155 segments was randomly selected for the non-incentive group from the pool of 2,064 segments in the national sample. The selected subsample included 1,838 residential addresses, and these households were expected to produce about 1,000 completed background questionnaires.

One adult was randomly selected in households with three or fewer eligible adults; two adults were selected in households with four or more eligible adults.

Only the impact of a $20 incentive was studied in the National Adult Literacy Survey. Since the field test did not indicate any important differences in response rates or data quality between a $20 and a $35 incentive, and the $20 incentive was more cost effective, the $35 incentive was eliminated and the research restricted to a comparison of a $20 incentive and no incentive.

Interviewers were assigned both non-incentive and incentive households in the areas they covered. This practice virtually eliminated interviewer effect as a factor in the differences between the two procedures. For the incentive households, interviewers were instructed to introduce the incentive before determining household composition and selecting the eligible respondent(s). Respondents who completed the background questionnaire and agreed to attempt the exercise booklet were given a $20 incentive check after completing the exercise booklet to the best of their ability.

As in the field test, three survey instruments were administered in the National Adult Literacy Survey: a screener, a background questionnaire, and an exercise booklet. The screener collected household-level data used in selecting the exercise respondents. The background questionnaire collected simple demographic data and detailed information on the selected respondent's education, labor force participation, and activities related to literacy. The exercise booklet was a 45-minute assessment of the respondent's prose, document, and quantitative literacy. The assessment was completed by the respondent, with instructions and timing guidance provided by the interviewer.

The interviewer introduced the incentive to the respondent before administering the screener. In the field test, respondents were informed of the incentive after the screener had been completed. As a result, incentives were expected to improve the screener response rate in the National Adult Literacy Survey, although there was no such effect in the field test.

An analysis of variance methodology was used to test the statistical significance of overall differences in response rates between the $0 and $20 incentive groups for the screener, the background questionnaire, and the exercise booklet . The formulas used to compute response rates were identical to those used for the field test (see section 10.2.3), except for differences in the exercise disposition coding schemes used in the 1991 field test and the 1992 National Adult Literacy Survey. However, the conceptual definition of the numerator and denominator of the exercise booklet response rates remained the same for the field test and the National Adult Literacy Survey.

10.3.3 Analysis of Response Rates

The field test provided information on response rates, data collection costs, the extent to which the sample was representative of the target population, and the intensity of respondents' efforts in the assessments. Consequently, this analysis concentrates on differences in response rates between the incentive and non-incentive subsamples.

The results of the analyses of the 1992 National Adult Literacy Survey data are consistent with those obtained in the 1991 field test. Response rates for both the background questionnaire and the exercise booklet improved substantially when respondents were offered incentives. Furthermore, as in the field test, the analysis showed that the incentives were most effective in increasing participation among persons with low levels of education and among minority populations who are generally underrepresented in national household surveys because of coverage problems in such surveys. As a result, the incentives improved the representation of the general population. Table 10-12 summarizes the response rates by incentive levels for the 1991 field test and the 1992 National Adult Literacy Survey. The sample design for the 1992 National Adult Literacy Survey provided for the oversampling of black and Hispanic adults to increase their sample size and thus to improve the ability to analyze the results for these minority groups. To account for the differential probabilities of selection of persons in the 1992 survey, weighted response rates were computed in addition to unweighted rates. The weights adjusted the sample representation of individuals to their representation in the total population. There are only small differences between weighted and unweighted response rates, as shown in Table 10-12 (see also Tables 10-14 and 10-15). A more detailed analysis of the results follows.

Table 10-12. Summary of response rates by incentive levels for the 1991 field test and the 1992 National Adult Literacy Survey

			1992 survey			
Incentive	1991 field test		$0 incentive		$20 incentive	
	$0	$20	Unweighted	Weighted[*]	Unweighted	Weighted[*]
Screener[†]	87.4%	87.7%	84.8%	–	88.8%	–
Background questionnaire	78.6%	82.7%	73.1%	71.8%	81.9%	81.1%
Exercise booklet	92.8%	97.9%	87.2%	89.2%	95.3%	95.8%
Overall	63.8%	71.0%	54.1%	54.3%	69.3%	69.0%

[*] The weighted response rates were calculated by applying a weight to each person to account for that person's probability of selection into the sample.

[†] For the 1991 field test, the incentive was introduced after the screener had been completed. For the 1992 survey, the incentive was introduced before the screener was completed.

10.3.3.1 Screener

Table 10-13 shows that the $20 incentive group's screener response rate was approximately 4 percentage points higher than that of the non-incentive cohort. This increase did not occur in the field test, because the field test respondents were not told about the incentives until the administration of the background questionnaire. Clearly, informing the potential respondents as early in the interview as possible enhances response rates.

Of the 24,827 households in the national sample, 3,164 units (13 percent) were either vacant or did not satisfy the definitions of a dwelling unit at the time of screening. As a result, 21,663 households were eligible to participate in the survey, of which 19,170 households (89 percent) completed the screener. Of 2,493 nonresponding households, 1,378 (55 percent) refused to participate in the study, and 1,115 (45 percent) did not complete the screener for other reasons.

Table 10-13 shows a surprising aspect of incentives. It was expected that the effect of incentives would be essentially restricted to a reduction in the screener refusal rate. However, incentives appear to have also reduced most other causes of screener nonresponse (i.e., respondents who were not at home despite repeated calls or some other reason).

Table 10-13. Screener response rates and reasons for nonresponse, by incentive condition: The 1992 National Adult Literacy Survey

Screener	$0 incentive			$20 incentive		
	Sample size	Percent of total	Percent of occupied units	Sample size	Percent of total	Percent of occupied units
Sampled dwelling units	1,838	100.0		22,989	100.0	
Occupied dwelling units	1,629	88.6	100.0	20,034	87.2	100.0
Completed screener	1,382	75.2	84.8	17,788	77.4	88.8
Refusal/breakoff	136	7.4	8.3	1,242	5.4	6.2
Language problem	12	0.6	0.7	68	0.3	0.3
Not home after max. attempts	55	3.0	3.4	575	2.5	2.8
Other nonresponse[*]	44	2.4	2.7	361	1.6	1.8
Response rate		84.8			88.8	

[*] Illness or disability, unavailable for field period, other.

10.3.3.2 Background questionnaire and exercise booklet

Table 10-14 presents the number of respondents who were eligible to complete the background questionnaire and the exercise booklet. A total of 17,863 respondents were eligible for the background questionnaire, of whom 81 percent provided completed questionnaires. A total of 3,346 did not complete this instrument, of whom 2,012 were refusals, 357 had some type of mental or physical disability that prevented their participation, 105 could not speak English or Spanish well enough to be interviewed and a suitable translator was not available, and the remainder were nonrespondents for other reasons (e.g., they were not at home despite repeated calls or were unavailable during the field period).

All 14,517 persons who completed the background questionnaire were asked to complete the exercise booklet. The exercise booklet response rate was 93 percent. Refusals were the dominant reason for nonresponse to the screener and the background questionnaire, but they accounted for only a small part of the nonresponse to the exercise booklet; most nonresponse was due to the respondent's inability to complete the exercises because of a language barrier, physical or mental disability, reading or writing barrier, or a similar factor. Apparently, the great majority of sampled persons who agreed to complete the background questionnaire were also willing to attempt the exercise booklet.

Table 10-14. Response rates and reasons for nonresponse for the background questionnaire and exercise booklet, by incentive condition: The 1992 National Adult Literacy Survey

	$0 incentive		$20 incentive	
	Sample size	Percent of total	Sample size	Percent of total
Background questionnaire				
Eligible respondents	1,273	100.0	16,590	100.0
Completed interview	930	73.1	13,587	81.9
Partial complete/breakoff	2	0.1	31	0.2
Refusal	211	16.6	1801	10.9
Language problem	5	0.4	100	0.6
Physical/mental disability	37	2.9	320	1.9
Not home after max. attempts	56	4.4	462	2.8
Other nonresponse[*]	32	2.5	289	1.7
Response rate		73.1		81.9
Weighted response rate		71.8		81.1
Exercise booklet				
Eligible respondents	930	100.0	13,591	100.0
Completed exercises	695	74.7	11,630	85.6
Partially completed exercises	82	8.8	1,029	7.6
Language barrier	7	0.7	223	1.7
Physical/mental disability	23	2.5	232	1.7
Reading/writing barrier	18	1.9	316	2.3
Breakoff, unwilling	11	1.2	79	0.6
Other nonresponse	23	2.5	179	1.3
Exercise not started	153	16.5	932	6.8
Language barrier	41	4.4	343	2.5
Physical/mental disability	26	2.8	191	1.4
Reading/writing barrier	18	2.0	79	0.6
Refusal	64	6.9	290	2.1
Other nonresponse	4	0.4	29	0.2
Response rate		87.2		95.3
Weighted response rate		89.2		95.8

[*] Illness or disability, unavailable for field period, other.

[+] Unavailable for field period, other.

Tables 10-13 and 10-14 show some surprising aspects of incentives. It was expected that the effect of incentives would be essentially restricted to a reduction in the refusal rate. However, in all three phases of the survey, incentives appear to have also reduced most other causes of nonresponse (in background question respondents who were not at home despite repeated calls and, physical and mental disability). It is not clear whether this reduction in nonresponse occurred because incentives made respondents try harder, had an effect on interviewer performance, or affected the reasons interviewers gave for nonresponse. In the field test, a similar pattern was seen for the background questionnaire but not for the exercise booklet.

Table 10-15. Response rates for the background questionnaire, by respondent characteristics and incentive condition: The 1992 National Adult Literacy Survey

	No incentive		$20 incentive	
Background questionnaire	Unweighted	Weighted*	Unweighted	Weighted*
All respondents	73.1	71.8	81.9	81.1
Age				
16-24	67.5	64.2	86.5	86.5
25-44	76.0	75.6	84.4	83.5
45-64	73.8	71.9	80.4	79.2
65+	71.5	70.0	77.4	74.9
Sex				
Male	70.7	70.1	79.6	78.7
Female	75.1	73.0	83.8	83.1
Race/ethnicity				
Hispanic	79.1	79.0	81.6	82.8
Black, non-Hispanic	72.1	70.6	85.4	84.9
White, non-Hispanic and other	72.6	71.4	80.9	80.5

*The weighted response rates were calculated by applying a weight to each individual to account for that individual's probability of selection into the sample.

A review of 1992 survey results for the background questionnaire in Table 10-15 and the exercise booklet in Table 10-16 basically confirms the findings of the 1991 field test. black adults demonstrated the greatest improvement in response as a result of incentives. The response rate for Hispanic adults was quite high, even without an incentive, and there was therefore not as much room for improvement.

Response rates rose dramatically for persons with relatively low levels of educational attainment, but there was only a small increase for college graduates and those with some college education. Incentives had a greater effect on young adults than on those age 45 or older, although the pattern was not as consistent as in the field test (in order from low to high).

Table 10-16. Response rates for the exercise booklet, by respondent characteristics and incentive condition: The 1992 National Adult Literacy Survey

	No incentive		$20 incentive	
Exercise booklet	Unweighted	Weighted*	Unweighted	Weighted*
All respondents	87.2	89.2	95.3	95.8
Age				
16-24	93.0	94.8	98.4	98.9
25-44	88.1	90.8	96.3	96.7
45-64	87.3	94.8	94.5	89.1
65+	79.8	79.2	88.9	90.6
Sex				
Male	86.6	88.2	95.1	95.6
Female	87.6	89.9	95.4	96.0
Race/ethnicity				
Hispanic	91.9	93.8	95.1	95.9
Black, non-Hispanic	79.1	82.2	94.1	94.5
White and other	88.9	89.7	95.7	96.0
Education level				
Some or no high school	81.5	84.4	92.5	93.1
High school graduate/GED	83.0	85.5	95.1	95.5
Some college or vocational education	91.7	93.1	96.8	97.3
College graduate or advanced degree	93.6	93.2	96.4	96.8

*The weighted response rates were calculated by applying a weight to each individual to account for that individual's probability of selection into the sample.

10.4 SUMMARY AND CONCLUSION

Two incentive experiments were conducted as part of the National Adult Literacy Survey design to assess the impact of various levels of incentives on the quality of the survey data. The first experiment, conducted as part of the 1991 field test, used three levels of incentives ($0, $20, and $35), which were randomly distributed among selected households.

The objective of the 1991 field test experiment was to ascertain the effect of incentives on test performance, response rates, and costs of the survey. The experiment indicated that (1) both a $20 and a $35 incentive significantly increased the response rate, especially for subgroups of particular interest to the study, but that $35 showed only a marginal improvement over $20; (2) the likelihood of producing a biased estimate of the population's literacy levels increases when no incentives are offered (i.e., due to the greater tendency for self-selection factors to occur in the $0 incentive group); and (3) the overall cost of conducting the assessment is less for the $20 incentive group than for the $0 and $35 incentive groups.

In the second experiment, conducted as part of the 1992 survey, two levels of incentives ($0 and $20) were randomly assigned to the national sample households. The analysis focused on the effect of incentives on response rates as a criterion for comparing the results of the 1991 field test and the 1992 survey.

The outcomes of the 1992 survey incentive experiment were similar to those of the 1991 field test study. The analysis of the incentive experiment confirmed that a $20 incentive increased the response rate. In fact, the 1992 survey showed a greater improvement than in the field test, but about half the improvement was due to the introduction of the incentive at the screener stage in the main study rather than at the background interview stage, which was the procedure in the field test. As indicated earlier, increasing the response rates for groups that are less likely to participate in the survey improves the distribution of the sample in its representation of the target population.

Chapter 11

EVALUATION OF SAMPLE DESIGN AND COMPOSITE ESTIMATION

John Burke, Leyla Mohadjer, James Green, Joseph Waksberg, Westat, Inc.

The National Adult Literacy Survey used a multistage sample design, as described in Chapter 2. The primary sampling units (PSUs) chosen at the first stage were metropolitan statistical areas (MSAs) or groups of non-MSA counties. The second-stage sampling units, referred to as segments, were individual census blocks or combinations of blocks. The third stage was a sample of dwelling units from listings made within the sampled segments. The final stage was the selection of individual household members within the selected dwelling units. In addition, the national sample design involved oversampling black and Hispanic adults.

Standard statistical methods of optimal allocation were used to arrive at approximations of an optimum design, when survey costs and estimates of intraclass correlations among selected items at each stage were taken into account. A common practice is to base design decisions on variances of important summary statistics derived from previous cycles of the same survey or similar surveys. The only comparable survey was the 1985 Young Adult Literacy Assessment, which included only persons ages 21-25 in the sample. Because only rough estimates of some of the variances were available from the 1985 survey, considerable judgment was involved in the development of the National Adult Literacy Survey design. It was therefore useful to evaluate the survey design after the survey had been completed. The sample design evaluation focused on the national component, because it is the component most certain to be repeated in a future survey. The results of this evaluation can be used in the design of future cycles of the National Adult Literacy Survey. Section 11.2 evaluates the National Adult Literacy Survey design by examining the components of variance for a selected group of literacy statistics.

One of the special features of the survey design was the application of composite estimation to produce weights for both national and state estimates. Data from the national sample were combined with the state data through the use of what are referred to as composite weights. A key objective of the national sample design was to provide reliable statistics for the adult population in the United States, as well as for some prespecified subdomains of interest (i.e., regional estimates and data for several race/ethnicity subgroups). Whereas the national sample was designed to provide reliable estimates at the national level, not every state was selected into the sample; furthermore, the sample sizes for states in the national sample were in most cases too small to produce reliable estimates for the individual states. To provide states with reliable literacy statistics, each state was invited to support a concurrent adult literacy survey with a sample design similar to the one used in the national survey. Samples for the 11 states participating in the concurrent state surveys were designed to be independent of the national sample (and independent among the states).

The national and state samples applied the same sampling procedures in terms of stratification method, PSU construction, sample design, and selection of various stages of sampling (except that black and Hispanic persons were oversampled for the national sample but not the state samples). Furthermore, the state and national surveys used the same instruments to screen households, collect background information, and conduct literacy assessments. To take full advantage of this comparability, the samples were combined to produce state-and national-level statistics. The advantage of compositing the samples was the increased sample size, which resulted in improved precision for both state and national estimates. It should be noted that composite estimates apply only to persons ages 16-64, because data for persons age 65 and over came only from the national sample.

The standard theoretical foundation of composite estimation requires a knowledge of variances of the statistics of interest. This information is necessary to produce the parameters used to combine data from various surveys in a way that minimizes the variances of the composite estimates. However, the composite weighting had to be completed before literacy score data were available. Section 11.2.4 describes the methods used to derive the compositing factors, which appeared to be reasonable approximations of the factors that would minimize variances in the National Adult Literacy Survey data. After the literacy data became available, new compositing factors were computed for a selected set of statistics. Section 11.3 describes the methodology used to derive the compositing factors, presents the estimated optimum compositing factors based on the National Adult Literacy Survey data, and provides an approach to be used for computing efficient compositing factors for the next cycle of the National Adult Literacy Survey.

The national sample design evaluation, and the compositing analysis, involved the computation of components of variance for a set of statistics chosen from the National Adult Literacy Survey data. The following section describes the statistics, the components of variance, and the statistical methods used to derive the outcome estimates.

11.1 ESTIMATION METHODS

Estimates of variances, design effects, and compositing factors were computed from the National Adult Literacy Survey data for (1) mean proficiency scores and (2) the percentage of persons scoring at each of five literacy levels: Level 1 (scores of 225 or lower), Level 2 (226 to 275), Level 3 (276 to 325), Level 4 (326 to 375), and Level 5 (376 or higher).

Estimates were computed for the following population totals:

1. Total population
2. Sex
3. Census region (northeast, midwest, south, and west)

4. Race/ethnicity (Hispanic; black, non-Hispanic; and other)
5. Education (less than high school diploma, high school diploma, some college, and college graduate)
6. Age (16-24, 25-44, 45-64, and 65 or older)
7. Country of birth (born in or outside of the United States)

Literacy levels were categorical data and the mean proficiency score was a continuous variable. These variables were expected to provide a useful indication of the range of the variability of estimates from the National Adult Literacy Survey data, with attention focused on particularly important statistics. Cases with missing demographic data were excluded from the variance calculations.

For a given population total, the usual unbiased weighted estimator is defined by

$$y' = \sum_{i=1}^{n} \frac{y_i}{\pi_i} \tag{1}$$

where

y' = the unbiased estimate of the population total Y;
n = the sample size;
y_i = the reported value of the characteristic for the i^{th} person in the sample; and
π_i = the probability of selection for the i^{th} respondent.

The variance of the estimate can be written as the sum of three components:

$$\sigma^2(y') = \sigma^2_{PSU} + \sigma^2_{Seg(PSU)} + \sigma^2_{HH(Seg)} \tag{2}$$

where

σ^2_{PSU} = the total contribution to variance resulting from the sampling of PSUs;
$\sigma^2_{Seg(PSU)}$ = the total contribution to variance resulting from the sampling of segments within PSUs; and
$\sigma^2_{HH(Seg)}$ = the total contribution to variance resulting from the sampling of households within segments.

Variances of estimates of means and ratios can also be decomposed in the manner shown above.

The variance of the unbiased estimate of a population total is given by equation (2). The term σ^2_{PSU} in equation (2) represents the contribution to variance resulting from the sampling of PSUs (in this case, counties or groups of counties). The value of σ^2_{PSU} depends in part on the number of PSUs in the sample. Specifically, if PCR is the number of certainty PSUs selected into the national sample, PNCR is the number of noncertainty PSUs, and P is the total number of PSUs (equal to PCR + PNCR), the between-PSU contribution to variance may be written approximately as

$$\sigma_{PSU}^2 = \frac{\sigma_I^2}{PNCR} \tag{3}$$

where σ_I^2 is the between-PSU unit variance among the noncertainty PSUs in the population. For the national sample, P = 101 and PNCR = 76.

The second and third terms in equation (2) reflect the total within-PSU variance resulting from the selection of segments within PSUs and of households within segments. If more than one household member was included in the sample, there would also be a within-household contribution to variance. If only one person is selected in a household, variance is there but it cannot be estimated unless there are two people selected.

The variance can be expressed in the following summarized form:

$$\sigma^2(y') = \sigma_B^2(y') + \sigma_W^2(y') \tag{4}$$

where

$\sigma^2 (y')$ = the total variance of the estimate;

$\sigma_B^2 (y')$ = the between-PSU component of variance; and

$\sigma_W^2 (y')$ = the within-PSU variance.

As mentioned above, the between-PSU component of variance reflects the contribution to variance that results from the sampling of PSUs. The within-PSU component reflects variability arising from several sources, including variance resulting from the selection of segments within PSUs, the selection of households within segments, and the selection of more than one person per household. This component also reflects the additional variability arising from the variation in weights due to the oversampling of black and Hispanic adults in segments with high concentrations of these minorities and the subsampling of persons within households.

11.1.1 Calculating Within- and Between-PSU Variance

Estimates of the components of variance were computed using the jackknife replication method. Under this approach, a set of replicates is formed where each replicate is a subset of the full sample. The replicate samples were formed by grouping all respondents by stratum and then randomly selecting a half-sample from one stratum. That half-sample was given a double weight. The process was repeated for other strata until the desired number of replicates was obtained. Each replicate provides an estimate of the statistic of interest, and the variability among the replicate estimates can be used to derive an estimate of the variance of the statistic (see Wolter, 1985). See Section 11.3 for a more detailed description of the jackknife method of estimating variances.

Depending on how the strata and pairs within strata are defined, the replication technique can also be used to estimate the separate components of variance shown in equation (3). For example, to estimate the total variance, $\sigma^2(y')$, the assignment of units within a stratum was made by pairing PSUs in noncertainty strata and pairing segments in certainty strata (see Section 3.3.1). Segments were placed in the original order of selection and assigned to each member of the pair in an alternate way. To estimate the within-PSU variances, $\sigma_W^2(y')$, the pairing was performed by segment in all strata, in both certainty and noncertainty PSUs.

The between-PSU variance was computed by subtraction as

$$\sigma_B^2(y') = \sigma^2(y') - \sigma_W^2(y') \tag{5}$$

11.2 THE NATIONAL SAMPLE DESIGN EVALUATION

Because the national survey used a multistage area probability sample, estimates from the survey are subject to larger sampling variances than would be expected from a simple random sample of the same size. This increased variance arises from the clustering of the sample by PSU, segments within PSUs, and households within segments and from the oversampling of black and Hispanic respondents in high-minority segments (i.e., those in which black and/or Hispanic adults accounted for 25 percent or more of the total population). There was also a small component in variance arising from the few households in which more than one person was included in the sample. This section provides estimates of the components of variance, design effects, and intraclass correlations arising from the national design.

The subset of the National Adult Literacy Survey data set in the national sample of PSUs was used to evaluate the national sample design. Because the original sample weights were computed for the national and state data combined, the weights for the national data set alone do not sum to the full population. Therefore, the weighting steps used in the survey (excluding the compositing step) were repeated on the national data set so that the weighted totals would be consistent with known population totals. The procedures used to compute final weights for the national data set were similar to those used for the combined the National Adult Literacy Survey data set (see Section 3.2.5). Furthermore, two sets of replicate weights were created for the national data set to permit the estimation of the total and within-PSU components of variance. The replicate weights for the total and within-PSU variances conformed to the description provided in the previous section.

Table 11-1 presents, for major population groups, estimates of four variance components (between- and within-PSU components, both with and without including variance due to imputation of proficiency) for each of six measures of prose literacy (the mean and the percentage in each of five levels). Tables 11-2 and 11-3 present the corresponding components of variance for document and quantitative literacy. The variance estimates are subject to fairly high sampling errors, and should be treated with caution. In some cases they are negative. The negative values shown in Tables 11-1 through 11-3 probably reflect the variances of the estimates of variance, which could show negative values when the true between-PSU components are small. Variances of components of variance tend to be relatively large in most surveys, and the numbers in these tables should probably not be taken as literal values; a smoothing process is useful to introduce more stability. Estimates of components of variance are given for the total population and for six major demographic subgroups, as specified in Section 11.1.

Table 11-1. Prose literacy variances of means and in percentages in five levels, both between and within PSU components, and both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations		Imputation variance incorporated?	Pcnt in Level 1 (225 or lower)		Pcnt in Level 2 (226 to 275)		Pcnt in Level 3 (276 to 325)		Pcnt in Lev 4 (326 to 375)		Pcnt in Lev 5 (376 or higher)		Average proficiency	
			Between	Within	Between	Within	Between	Within	Between	Within	Between	Within	Between	Within
Total Population														
	Total	No	0.08	0.10	0.09	0.26	0.03	0.21	0.02	0.12	0.00	0.03	0.21	0.17
		Yes	0.08	0.13	0.09	0.32	0.03	0.43	0.02	0.19	0.00	0.09	0.21	0.19
Sex														
	Male	No	0.05	0.37	-0.02	0.53	0.00	0.49	-0.05	0.36	-0.01	0.08	-0.01	0.85
		Yes	0.05	0.40	-0.02	0.83	0.00	1.06	-0.05	0.58	-0.01	0.17	-0.01	0.91
	Female	No	0.09	0.18	0.05	0.35	0.07	0.44	0.04	0.23	-0.02	0.06	0.30	0.37
		Yes	0.09	0.25	0.05	0.49	0.07	0.95	0.04	0.31	-0.02	0.11	0.30	0.50
Census Region														
	Northeast	No	0.03	1.35	-0.25	2.19	0.64	2.14	-0.03	0.46	0.03	0.09	-0.73	3.51
		Yes	0.03	1.50	-0.25	2.81	0.64	2.39	-0.03	1.08	0.03	0.14	-0.73	3.77
	Midwest	No	0.21	0.55	0.49	0.88	0.45	0.50	-0.10	0.62	-0.05	0.14	1.00	1.02
		Yes	0.21	0.89	0.49	1.04	0.45	1.25	-0.10	0.79	-0.05	0.28	1.00	1.05
	South	No	0.36	0.47	0.20	1.03	-0.11	0.64	0.48	0.36	0.01	0.07	1.65	0.99
		Yes	0.36	0.63	0.20	1.23	-0.11	0.98	0.48	0.77	0.01	0.10	1.65	1.35
	West	No	-0.06	0.81	0.13	1.26	-0.23	0.88	-0.28	1.34	-0.02	0.34	0.03	2.61
		Yes	-0.06	1.42	0.13	2.22	-0.23	1.71	-0.28	2.36	-0.02	0.68	0.03	2.96
Race/Ethnicity														
	Black	No	-0.48	1.51	-0.35	1.52	0.26	0.74	0.02	0.09	0.00	0.02	-0.20	2.07
		Yes	-0.48	2.11	-0.35	1.75	0.26	0.93	0.02	0.30	0.00	0.04	-0.20	2.50
	Hispanic	No	1.39	2.09	0.54	2.06	0.74	0.80	-0.08	0.56	-0.01	0.08	2.23	5.02
		Yes	1.39	2.44	0.54	2.92	0.74	2.17	-0.08	1.18	-0.01	0.15	2.23	5.21
	Other	No	0.12	0.13	0.12	0.34	0.01	0.33	0.03	0.18	0.00	0.05	0.36	0.27
		Yes	0.12	0.17	0.12	0.39	0.01	0.56	0.03	0.27	0.00	0.16	0.36	0.30
Education Level														
	No HS degree	No	-0.05	1.07	-0.19	1.07	-0.02	0.53	0.00	0.12	0.00	0.00	0.28	1.47
		Yes	-0.05	1.34	-0.19	1.52	-0.02	0.98	0.00	0.22	0.00	0.01	0.28	1.83
	HS degree	No	0.25	0.50	0.49	0.97	0.36	0.69	0.03	0.39	0.00	0.02	0.53	0.66
		Yes	0.25	0.72	0.49	1.18	0.36	1.88	0.03	0.63	0.00	0.06	0.53	0.91
	Some college	No	0.07	0.17	0.19	0.53	0.32	0.78	0.14	0.68	-0.02	0.16	0.36	0.86
		Yes	0.07	0.26	0.19	0.63	0.32	0.82	0.14	0.75	-0.02	0.18	0.36	1.13
	College graduate	No	0.01	0.22	0.04	0.42	-0.22	1.08	-0.11	1.17	-0.11	0.61	-0.02	1.36
		Yes	0.01	0.31	0.04	0.81	-0.22	1.56	-0.11	2.26	-0.11	1.40	-0.02	1.96
Age														
	16 - 24 years	No	-0.30	0.88	0.14	1.76	0.38	1.20	0.14	0.76	-0.01	0.07	0.70	0.99
		Yes	-0.30	1.19	0.14	3.09	0.38	2.02	0.14	1.37	-0.01	0.11	0.70	1.03
	25 - 44 years	No	0.06	0.24	-0.15	0.60	-0.11	0.66	-0.01	0.43	0.01	0.09	0.10	0.70
		Yes	0.06	0.34	-0.15	0.66	-0.11	0.77	-0.01	0.54	0.01	0.20	0.10	0.76
	45 - 64 years	No	0.36	0.37	0.04	0.75	-0.11	1.03	-0.03	0.55	-0.02	0.15	0.83	0.88
		Yes	0.36	0.45	0.04	1.03	-0.11	1.40	-0.03	0.79	-0.02	0.23	0.83	0.92
	65 years and older	No	1.00	1.04	0.70	1.21	0.17	0.85	-0.04	0.32	0.00	0.04	1.36	1.93
		Yes	1.00	1.56	0.70	1.52	0.17	1.49	-0.04	0.81	0.00	0.13	1.36	3.06
Country of Birth														
	Not USA	No	0.35	1.81	-0.09	1.89	0.22	1.11	-0.12	0.78	0.00	0.06	1.27	4.49
		Yes	0.35	2.23	-0.09	1.93	0.22	1.76	-0.12	0.88	0.00	0.23	1.27	4.97
	USA	No	0.10	0.09	0.13	0.28	0.01	0.27	0.03	0.16	0.00	0.04	0.26	0.22
		Yes	0.10	0.13	0.13	0.34	0.01	0.52	0.03	0.26	0.00	0.11	0.26	0.27

Table 11-2. Document literacy variances of means and in percentages in five levels, both between and within PSU components, and both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Pcnt in Level 1 (225 or lower)		Pcnt in Level 2 (226 to 275)		Pcnt in Level 3 (276 to 325)		Pcnt in Level 4 (326 to 375)		Pcnt in Level 5 (376 or higher)		Average proficiency	
		Between	Within	Between	Within	Between	Within	Between	Within	Between	Within	Between	Within
Total Population													
Total	No	0.13	0.10	0.07	0.15	-0.02	0.27	0.03	0.11	0.00	0.03	0.26	0.20
	Yes	0.13	0.12	0.07	0.30	-0.02	0.52	0.03	0.19	0.00	0.07	0.26	0.25
Sex													
Male	No	-0.04	0.47	0.05	0.32	-0.03	0.45	-0.05	0.36	-0.01	0.07	-0.06	1.03
	Yes	-0.04	0.51	0.05	0.46	-0.03	1.29	-0.05	0.58	-0.01	0.10	-0.06	1.08
Female	No	0.22	0.18	-0.08	0.38	-0.02	0.46	0.03	0.20	0.00	0.03	0.50	0.31
	Yes	0.22	0.24	-0.08	0.64	-0.02	0.54	0.03	0.42	0.00	0.09	0.50	0.39
Census Region													
Northeast	No	-0.20	1.41	0.08	1.49	-0.24	2.58	-0.28	0.73	-0.01	0.07	-1.42	4.37
	Yes	-0.20	1.84	0.08	2.57	-0.24	3.54	-0.28	1.69	-0.01	0.13	-1.42	4.64
Midwest	No	0.43	0.63	0.30	0.56	-0.19	0.91	0.32	0.70	0.00	0.05	2.11	1.36
	Yes	0.43	0.69	0.30	1.19	-0.19	3.32	0.32	1.08	0.00	0.15	2.11	1.44
South	No	0.87	0.38	0.06	0.49	0.12	0.53	0.27	0.23	0.01	0.09	1.98	1.14
	Yes	0.87	0.57	0.06	0.72	0.12	1.00	0.27	0.25	0.01	0.11	1.98	1.25
West	No	-0.19	1.07	0.49	0.39	-0.01	1.29	0.29	1.00	-0.06	0.29	-0.30	2.33
	Yes	-0.19	1.16	0.49	1.72	-0.01	1.98	0.29	2.08	-0.06	0.46	-0.30	2.71
Race/Ethnicity													
Black	No	0.02	1.17	-0.12	1.20	0.26	0.81	-0.01	0.12	0.00	0.02	-0.21	2.20
	Yes	0.02	1.53	-0.12	1.67	0.26	0.88	-0.01	0.15	0.00	0.02	-0.21	2.24
Hispanic	No	1.67	1.76	0.68	1.04	0.56	0.98	-0.03	0.63	0.00	0.04	2.89	5.77
	Yes	1.67	2.11	0.68	2.67	0.56	1.78	-0.03	0.80	0.00	0.16	2.89	5.84
Other	No	0.18	0.13	0.05	0.20	-0.05	0.46	0.03	0.19	0.00	0.04	0.46	0.29
	Yes	0.18	0.19	0.05	0.32	-0.05	0.82	0.03	0.29	0.00	0.11	0.46	0.35
Education Level													
No HS degree	No	0.41	0.82	0.39	0.75	-0.06	0.63	0.00	0.13	0.00	0.01	0.58	1.57
	Yes	0.41	1.30	0.39	1.35	-0.06	1.47	0.00	0.17	0.00	0.01	0.58	1.89
HS degree	No	0.36	0.46	-0.03	0.71	0.11	0.62	0.06	0.29	0.00	0.02	0.63	0.67
	Yes	0.36	0.64	-0.03	0.84	0.11	1.39	0.06	0.45	0.00	0.04	0.63	0.77
Some college	No	0.07	0.21	0.00	0.68	-0.03	1.06	0.24	0.49	0.02	0.08	0.50	0.81
	Yes	0.07	0.23	0.00	0.80	-0.03	2.02	0.24	0.82	0.02	0.12	0.50	0.84
College graduate	No	-0.01	0.16	0.07	0.54	-0.25	1.05	0.18	1.35	-0.06	0.41	0.09	1.22
	Yes	-0.01	0.26	0.07	0.93	-0.25	1.37	0.18	2.32	-0.06	1.30	0.09	1.45
Age													
16 – 24 years	No	-0.11	0.64	0.49	0.95	0.43	1.02	0.09	0.65	0.01	0.08	0.30	1.29
	Yes	-0.11	0.95	0.49	2.65	0.43	6.59	0.09	2.71	0.01	0.16	0.30	1.42
25 – 44 years	No	0.18	0.20	0.03	0.35	-0.17	0.72	-0.07	0.52	-0.01	0.09	0.16	0.82
	Yes	0.18	0.23	0.03	0.47	-0.17	1.13	-0.07	0.73	-0.01	0.21	0.16	0.88
45 – 64 years	No	0.56	0.39	-0.01	0.81	-0.24	1.06	0.21	0.34	0.01	0.08	0.74	0.88
	Yes	0.56	0.74	-0.01	1.30	-0.24	1.94	0.21	0.54	0.01	0.23	0.74	1.34
65 years and older	No	0.76	1.30	0.33	1.34	0.11	0.65	-0.04	0.15	0.00	0.01	1.29	1.82
	Yes	0.76	1.39	0.33	1.43	0.11	0.99	-0.04	0.34	0.00	0.02	1.29	2.48
Country of Birth													
Not USA	No	0.31	1.86	-0.08	1.22	-0.08	1.13	-0.08	0.79	0.01	0.05	0.66	5.29
	Yes	0.31	2.58	-0.08	2.20	-0.08	1.31	-0.08	1.16	0.01	0.14	0.66	5.73
USA	No	0.14	0.10	0.06	0.18	-0.04	0.34	0.04	0.13	0.00	0.04	0.33	0.22
	Yes	0.14	0.13	0.06	0.32	-0.04	0.64	0.04	0.22	0.00	0.08	0.33	0.27

Table 11-3. Quantitative literacy variances of means and in percentages in each of five levels, both between and within PSU components, with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Pcnt in Level 1 (225 or lower)		Pcnt in Level 2 (226 to 275)		Pcnt in Level 3 (276 to 325)		Pcnt in Level 4 (326 to 375)		Pcnt in Level 5 (376 or higher)		Average proficiency	
		Between	Within	Between	Within	Between	Within	Between	Within	Between	Within	Between	Within
Total Population													
Total	No	0.10	0.09	-0.01	0.20	-0.04	0.30	0.01	0.16	-0.00	0.03	0.29	0.19
	Yes	0.10	0.17	-0.01	0.28	-0.04	0.37	0.01	0.24	-0.00	0.05	0.29	0.30
Sex													
Male	No	0.12	0.27	-0.02	0.36	-0.18	0.62	0.01	0.37	-0.00	0.13	0.07	0.88
	Yes	0.12	0.38	-0.02	0.45	-0.18	0.90	0.01	0.66	-0.00	0.17	0.07	1.02
Female	No	0.01	0.28	-0.03	0.47	0.04	0.49	-0.00	0.25	-0.01	0.05	0.21	0.58
	Yes	0.01	0.35	-0.03	0.98	0.04	1.10	-0.00	0.41	-0.01	0.09	0.21	0.73
Census Region													
Northeast	No	0.28	0.67	0.22	0.45	-0.40	1.34	-0.77	1.28	-0.04	0.22	-0.54	2.98
	Yes	0.28	1.48	0.22	2.14	-0.40	1.66	-0.77	1.56	-0.04	0.33	-0.54	3.71
Midwest	No	0.68	0.58	0.47	0.81	-0.25	1.14	0.34	0.81	-0.01	0.11	2.10	1.17
	Yes	0.68	0.81	0.47	2.04	-0.25	2.23	0.34	1.30	-0.01	0.27	2.10	1.43
South	No	0.29	0.52	0.01	0.62	0.33	0.63	0.29	0.25	-0.00	0.10	1.51	1.37
	Yes	0.29	0.66	0.01	1.18	0.33	0.87	0.29	0.45	-0.00	0.11	1.51	1.85
West	No	-0.08	0.73	-0.44	1.18	-0.73	1.95	-0.20	1.73	0.02	0.21	-0.36	3.26
	Yes	-0.08	0.94	-0.44	1.25	-0.73	2.24	-0.20	2.00	0.02	0.24	-0.36	3.32
Race/Ethnicity													
Black	No	0.19	1.46	-0.38	1.27	0.07	0.71	0.03	0.10	-0.00	0.01	0.56	2.15
	Yes	0.19	2.11	-0.38	2.07	0.07	1.00	0.03	0.16	-0.00	0.01	0.56	2.68
Hispanic	No	0.52	1.58	0.22	1.55	0.59	0.90	0.11	0.50	-0.01	0.03	1.89	6.05
	Yes	0.52	1.73	0.22	2.44	0.59	1.97	0.11	1.24	-0.01	0.05	1.89	6.73
Other	No	0.18	0.10	0.03	0.24	-0.02	0.46	0.04	0.23	-0.00	0.05	0.50	0.27
	Yes	0.18	0.21	0.03	0.32	-0.02	0.50	0.04	0.34	-0.00	0.08	0.50	0.47
Education Level													
No HS degree	No	0.49	0.81	-0.01	0.89	0.02	0.35	0.02	0.10	-0.00	0.02	1.02	1.57
	Yes	0.49	1.31	-0.01	1.88	0.02	0.67	0.02	0.21	-0.00	0.03	1.02	1.75
HS degree	No	0.33	0.36	0.23	0.86	0.06	0.94	-0.05	0.48	0.00	0.04	0.49	0.67
	Yes	0.33	0.49	0.23	1.25	0.06	1.06	-0.05	0.84	0.00	0.06	0.49	0.75
Some college	No	0.02	0.25	0.03	0.66	0.16	0.86	0.34	0.66	-0.02	0.15	0.46	1.01
	Yes	0.02	0.28	0.03	0.79	0.16	1.57	0.34	1.51	-0.02	0.23	0.46	1.22
College graduate	No	0.00	0.16	-0.15	0.50	-0.73	2.06	0.12	1.42	0.03	0.63	-0.15	1.51
	Yes	0.00	0.18	-0.15	0.67	-0.73	2.41	0.12	1.55	0.03	0.74	-0.15	2.14
Age													
16 – 24 years	No	0.10	0.68	0.13	1.05	-0.32	1.71	0.22	0.60	0.01	0.09	0.35	1.49
	Yes	0.10	0.86	0.13	1.40	-0.32	1.89	0.22	0.63	0.01	0.11	0.35	1.67
25 – 44 years	No	0.11	0.22	-0.04	0.47	-0.11	0.75	-0.16	0.60	0.01	0.11	0.22	0.70
	Yes	0.11	0.28	-0.04	0.53	-0.11	0.90	-0.16	0.88	0.01	0.23	0.22	0.76
45 – 64 years	No	0.45	0.40	-0.17	0.90	-0.24	1.27	0.29	0.58	-0.03	0.19	1.13	0.92
	Yes	0.45	0.73	-0.17	1.91	-0.24	2.03	0.29	0.87	-0.03	0.20	1.13	1.15
65 years and older	No	1.11	1.28	-0.16	1.28	0.19	1.14	0.00	0.33	0.03	0.14	3.55	2.32
	Yes	1.11	1.63	-0.16	1.61	0.19	1.46	0.00	0.43	0.03	0.19	3.55	3.30
Country of Birth													
Not USA	No	0.10	1.50	-0.51	1.54	-0.17	1.31	-0.24	1.05	0.01	0.09	0.54	6.80
	Yes	0.10	2.25	-0.51	2.23	-0.17	1.59	-0.24	1.51	0.01	0.27	0.54	7.07
USA	No	0.13	0.10	0.00	0.22	-0.02	0.34	0.03	0.17	-0.00	0.05	0.38	0.22
	Yes	0.13	0.16	0.00	0.28	-0.02	0.41	0.03	0.24	-0.00	0.07	0.38	0.32

A comparison of the magnitude of the between- and within-PSU variances shows that, in general, the between-PSU variance constitutes a small portion of the total variance. The importance of the between-PSU component of variance is that it determines a lower bound on the sampling errors associated with a multistage sample design. Increases in the numbers of sampled segments and households cannot reduce the variance below the level determined by the between-PSU variance. Consequently, useful reductions in variances can be achieved only by increasing the number of PSUs in the sample. The between-PSU variances are usually quite small compared to the within-PSU variances. The few large values are mostly statistics for the total population (e.g., mean overall proficiency for all persons) or for large population subgroups (e.g., mean overall proficiency for persons born in the United States or for those reporting race/ethnicity as "other"). It is likely that this is due to the large sample sizes per PSU for these items, as compared to the sizes for subgroup analysis.

The small values of the between-PSU variances for most items indicate that, without an increase in the number of sample cases, an increase in the number of PSUs would have very little effect on the variances for subgroup analysis and would thus be an inefficient design. Increasing the number of PSUs would usefully reduce variance for statistics relating to the total population, but such statistics have quite small variances anyway because of the large sample size. Further reductions are probably not necessary for the kinds of data analyses contemplated for the total population. The results indicate that the next cycle of the National Adult Literacy Survey should use approximately the same number of national sample PSUs as the 1992 survey.

11.2.1 Computing the Design Effect

For a particular subgroup, the estimated mean proficiency score, or the proportion of respondents at a given literacy level, can be expressed in the following way:

$$r' = \frac{y'}{x'} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i x_i} \qquad (6)$$

where

w_i = the sample weight attached to the i^{th} respondent;

n = the sample size; and

$$x_i = \begin{cases} 1 & \text{if the respondent belongs to the subgroup} \\ 0 & \text{otherwise} \end{cases}$$

If r' estimates the mean proficiency score,

$$y_i = \begin{cases} \text{the proficiency score if the } i^{th} \text{ respondent belongs to the subgroup} \\ 0 \text{ if otherwise} \end{cases}$$

Otherwise, if r' estimates the proportion of persons at a given literacy level,

$$y_i = \begin{cases} 1 \text{ if the proficiency score for the } i^{th} \text{ respondent was at the level, and} \\ 0 \text{ otherwise} \end{cases}$$

The design effect (DEFF) was determined by calculating the ratio of variances defined by

$$\text{DEFF} = \frac{\sigma^2(r')}{\sigma^2_{SRS}} \tag{7}$$

where σ^2_{SRS} is the corresponding variance of r based on a simple random sample. The design effect provides an overall measure of the increase in variance associated with the use of a multistage sample instead of a simple random sample of the same size.

The jackknife replication technique was used to estimate the variance of r. The σ^2_{SRS} (r^2 cr)was estimated in the following way:

$$\sigma^2(\text{SRS}) = \frac{1}{n}\,\frac{1}{(x')^2}\left[\operatorname{var}(y') + r^2 \operatorname{var}(x') - 2r' \operatorname{cov}(y', x')\right] \tag{8}$$

The SAS procedure CORR was used to estimate the variances and covariances in equation (8).

Tables 11-4, 11-5, and 11-6 show the estimated design effects for the national data on selected estimates for the prose, document, and quantitative scales, respectively. Design effects are given for estimates of proportions and mean proficiencies for the total population and for six major demographic subgroups, as specified in Section 11.1. The design effects are shown in the tables both with and without the imputation variance resulting from the use of plausible values in estimation (see Section 9.4). It is important to note that the design effects provided in these tables are for the national sample design only (i.e., they do not include any of the state data) and should not be used as estimates of overall design effects for the National Adult Literacy Survey. As the tables show, the design effects are relatively low for most of the statistics. The design effect, including the imputation variance, was computed as

$$\text{DEFF} = \frac{\sigma^2(r') + \sigma^2_{imp}}{\sigma^2_{SRS} + \sigma^2_{imp}} \tag{9}$$

where σ^2_{imp} is the variance due to imputation.

Table 11-4. Prose literacy design effects for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	No	1.51	2.47	1.53	1.39	1.37	1.19
	Yes	1.41	2.06	1.23	1.23	1.10	1.18
Sex							
Male	No	1.64	1.76	1.46	1.39	1.34	1.17
	Yes	1.56	1.38	1.17	1.20	1.13	1.15
Female	No	1.20	1.41	1.63	1.34	1.01	1.17
	Yes	1.16	1.28	1.24	1.23	1.00	1.14
Census Region							
Northeast	No	2.33	2.75	3.58	0.92	1.42	1.86
	Yes	2.06	1.94	2.96	0.96	1.27	1.73
Midwest	No	1.86	2.21	1.33	1.14	0.90	1.95
	Yes	1.47	1.97	1.16	1.10	0.96	1.93
South	No	2.18	2.81	1.20	3.03	1.32	2.78
	Yes	1.84	2.25	1.11	1.82	1.20	2.29
West	No	1.30	2.37	0.86	1.77	2.39	1.41
	Yes	1.15	1.52	0.93	1.29	1.40	1.34
Race/Ethnicity							
Black	No	0.65	0.75	0.89	0.46	1.05	0.81
	Yes	0.74	0.78	0.91	0.71	1.03	0.84
Hispanic	No	1.85	1.83	1.40	1.16	1.53	1.51
	Yes	1.71	1.52	1.18	1.06	1.21	1.49
Other	No	2.20	2.61	1.58	1.43	1.32	1.95
	Yes	1.90	2.25	1.28	1.28	1.08	1.86
Education Level							
No HS degree	No	1.49	1.50	1.43	1.53	0.26	1.46
	Yes	1.35	1.29	1.19	1.24	0.98	1.35
HS degree	No	2.30	2.59	1.87	1.88	1.93	2.09
	Yes	1.77	2.17	1.28	1.42	1.23	1.76
Some college	No	1.33	1.36	1.48	1.54	1.75	1.84
	Yes	1.21	1.31	1.46	1.48	1.59	1.60
College graduate	No	1.90	1.43	0.97	1.06	1.11	1.55
	Yes	1.53	1.19	0.98	1.03	1.04	1.32
Age							
16 - 24 years	No	1.21	2.18	1.67	1.71	1.03	1.46
	Yes	1.13	1.47	1.36	1.33	1.02	1.44
25 - 44 years	No	1.33	1.48	1.40	1.38	1.18	1.07
	Yes	1.23	1.40	1.31	1.28	1.07	1.07
45 - 64 years	No	1.47	1.28	1.36	1.21	1.29	1.30
	Yes	1.40	1.19	1.23	1.13	1.16	1.29
65 years and older	No	1.78	1.87	1.45	1.08	1.38	1.59
	Yes	1.53	1.67	1.23	1.03	1.10	1.38
Country of Birth							
Not USA	No	1.16	1.39	1.25	1.40	0.86	1.06
	Yes	1.13	1.38	1.16	1.32	0.96	1.05
USA	No	1.67	2.49	1.53	1.53	1.46	1.64
	Yes	1.50	2.06	1.23	1.29	1.13	1.56

Table 11-5. Document literacy design effects for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	No	1.51	2.47	1.53	1.39	1.37	1.19
	Yes	1.41	2.06	1.23	1.23	1.10	1.18
Sex							
Male	No	1.64	1.76	1.46	1.39	1.34	1.17
	Yes	1.56	1.38	1.17	1.20	1.13	1.15
Female	No	1.20	1.41	1.63	1.34	1.01	1.17
	Yes	1.16	1.28	1.24	1.23	1.00	1.14
Census Region							
Northeast	No	2.33	2.75	3.58	0.92	1.42	1.86
	Yes	2.06	1.94	2.96	0.96	1.27	1.73
Midwest	No	1.86	2.21	1.33	1.14	0.90	1.95
	Yes	1.47	1.97	1.16	1.10	0.96	1.93
South	No	2.18	2.81	1.20	3.03	1.32	2.78
	Yes	1.84	2.25	1.11	1.82	1.20	2.29
West	No	1.30	2.37	0.86	1.77	2.39	1.41
	Yes	1.15	1.52	0.93	1.29	1.40	1.34
Race/Ethnicity							
Black	No	0.65	0.75	0.89	0.46	1.05	0.81
	Yes	0.74	0.78	0.91	0.71	1.03	0.84
Hispanic	No	1.85	1.83	1.40	1.16	1.53	1.51
	Yes	1.71	1.52	1.18	1.06	1.21	1.49
Other	No	2.20	2.61	1.58	1.43	1.32	1.95
	Yes	1.90	2.25	1.28	1.28	1.08	1.86
Education Level							
No HS degree	No	1.49	1.50	1.43	1.53	0.26	1.46
	Yes	1.35	1.29	1.19	1.24	0.98	1.35
HS degree	No	2.30	2.59	1.87	1.88	1.93	2.09
	Yes	1.77	2.17	1.28	1.42	1.23	1.76
Some college	No	1.33	1.36	1.48	1.54	1.75	1.84
	Yes	1.21	1.31	1.46	1.48	1.59	1.60
College graduate	No	1.90	1.43	0.97	1.06	1.11	1.55
	Yes	1.53	1.19	0.98	1.03	1.04	1.32
Age							
16 - 24 years	No	1.21	2.18	1.67	1.71	1.03	1.46
	Yes	1.13	1.47	1.36	1.33	1.02	1.44
25 - 44 years	No	1.33	1.48	1.40	1.38	1.18	1.07
	Yes	1.23	1.40	1.31	1.28	1.07	1.07
45 - 64 years	No	1.47	1.28	1.36	1.21	1.29	1.30
	Yes	1.40	1.19	1.23	1.13	1.16	1.29
65 years and older	No	1.78	1.87	1.45	1.08	1.38	1.59
	Yes	1.53	1.67	1.23	1.03	1.10	1.38
Country of Birth							
Not USA	No	1.16	1.39	1.25	1.40	0.86	1.06
	Yes	1.13	1.38	1.16	1.32	0.96	1.05
USA	No	1.67	2.49	1.53	1.53	1.46	1.64
	Yes	1.50	2.06	1.23	1.29	1.13	1.56

Table 11-6. Quantitative literacy design effects for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	No	1.51	1.42	1.65	1.64	1.11	1.31
	Yes	1.31	1.27	1.47	1.36	1.07	1.24
Sex							
Male	No	1.60	1.28	1.33	1.54	1.53	1.18
	Yes	1.41	1.21	1.18	1.25	1.33	1.16
Female	No	1.15	1.50	1.77	1.38	1.03	1.17
	Yes	1.12	1.18	1.25	1.21	1.01	1.14
Census Region							
Northeast	No	1.48	1.04	1.21	1.05	1.55	1.34
	Yes	1.21	1.01	1.15	1.03	1.28	1.25
Midwest	No	3.01	2.10	1.28	2.35	0.85	2.70
	Yes	2.30	1.36	1.11	1.67	0.94	2.40
South	No	2.02	1.49	2.19	1.96	1.20	2.63
	Yes	1.76	1.21	1.76	1.56	1.18	2.14
West	No	1.09	1.24	1.63	2.73	1.38	1.44
	Yes	1.07	1.22	1.45	2.18	1.34	1.42
Race/Ethnicity							
Black	No	0.99	0.59	0.83	0.64	0.74	0.96
	Yes	0.99	0.73	0.87	0.72	0.75	0.97
Hispanic	No	1.12	1.27	1.35	1.43	0.66	1.51
	Yes	1.11	1.17	1.18	1.15	0.79	1.45
Other	No	2.41	1.55	2.11	1.71	1.17	2.10
	Yes	1.73	1.37	1.95	1.40	1.11	1.70
Education Level							
No HS degree	No	1.91	1.59	1.07	1.22	1.78	1.76
	Yes	1.52	1.21	1.03	1.11	1.44	1.68
HS degree	No	1.94	2.00	1.78	1.94	1.50	1.78
	Yes	1.68	1.59	1.65	1.36	1.30	1.70
Some college	No	1.37	1.30	1.41	1.79	1.03	1.89
	Yes	1.32	1.24	1.21	1.32	1.02	1.70
College graduate	No	1.36	1.00	1.48	1.58	1.33	1.39
	Yes	1.31	1.00	1.35	1.51	1.27	1.24
Age							
16 - 24 years	No	1.35	1.38	1.50	1.65	1.45	1.46
	Yes	1.27	1.27	1.42	1.61	1.37	1.40
25 - 44 years	No	1.37	1.42	1.64	1.44	1.27	1.16
	Yes	1.29	1.34	1.46	1.23	1.12	1.15
45 - 64 years	No	1.73	1.20	1.55	1.91	1.33	1.40
	Yes	1.44	1.08	1.26	1.55	1.30	1.34
65 years and older	No	2.06	1.26	1.81	1.03	1.75	1.94
	Yes	1.81	1.19	1.56	1.03	1.50	1.71
Country of Birth							
Not USA	No	0.86	0.77	0.99	1.39	1.25	1.19
	Yes	0.90	0.85	0.99	1.22	1.08	1.18
USA	No	1.84	1.41	1.79	1.68	1.25	1.72
	Yes	1.58	1.29	1.56	1.44	1.15	1.56

It is evident from the figures reported in Tables 11-4 to 11-6 that the design effect becomes closer to one as the imputation variance is added to the variances.

An equivalent expression for DEFF that is useful for examining the contributions of clustering and unequal probabilities of selection to the variances is given by

$$\mathrm{DEFF} = 1 + CV^2(w) + \rho(\bar{m} - 1) \qquad (10)$$

where

$CV^2(w) = \dfrac{\mathrm{var}(w)}{\bar{w}^2}$, the square of the coefficient of variation in the sample weights (relvariance);

$\bar{w}$ = the average value of the weights;

ρ = the within-PSU intraclass correlation; and

$\bar{m}$ = the average number of respondents per PSU.

Tables 11-7, 11-8 and 11-9 show the estimated within-PSU intraclass correlations for the prose, document, and quantitative literacy scales. The intraclass correlations were computed by solving for ρ in equation (10). Intraclass correlations are given for the variables of interest both without and with (as shown in parentheses) the imputation variance (see Chapter 9). As indicated in the tables, the within-PSU intraclass correlations are small overall. However, it must be noted that these intraclass correlations are, on average, larger than the 0.00075 assumed in the derivation of compositing factors, as described in Section 3.2.4.3.

Table 11-10 shows the increase in variance resulting from the variability of the sample weights, the term $1 + CV^2(w)$ in equation (9). Variability in weights was caused mainly by three factors: the oversampling of black and Hispanic respondents, the subsampling of persons within households, and the adjustment of weights through poststratification. A comparison of the design effects (Tables 11-4, 11-5, and 11-6) with the intraclass correlations (Tables 11-7, 11-8, and 11-9) and the relvariance of the weights (Table 11-10) shows that the main source of increase in the design effects is the variability of the sample weights. For nonminority groups, the variability in the weights comes mainly from the poststratification of the weights. Among the minority groups, a larger portion of the variability in the weights is attributable to oversampling than to poststratification. The intraclass correlations within PSUs are very small, suggesting that the effect of clustering was quite limited in the national design. If the correlations are small, one may be able to increase the number of segments per PSU and reduce the number of PSUs and get the same variance at lower cost.

Table 11-7. Prose literacy intraclass correlation for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	No	0.001	0.008	0.001	0.000	0.000	-0.001
	Yes	0.001	0.005	-0.001	-0.001	-0.002	-0.001
Sex							
Male	No	0.006	0.008	0.003	0.002	0.001	-0.002
	Yes	0.005	0.001	-0.002	-0.002	-0.003	-0.003
Female	No	-0.002	0.001	0.004	-0.000	-0.004	-0.002
	Yes	-0.002	-0.001	-0.001	-0.001	-0.004	-0.003
Census Region							
Northeast	No	0.008	0.012	0.019	-0.003	0.001	0.004
	Yes	0.006	0.005	0.013	-0.003	-0.000	0.003
Midwest	No	0.005	0.008	0.000	-0.001	-0.003	0.005
	Yes	0.001	0.006	-0.001	-0.002	-0.003	0.005
South	No	0.006	0.010	-0.001	0.011	-0.000	0.009
	Yes	0.003	0.006	-0.001	0.003	-0.001	0.006
West	No	-0.000	0.008	-0.004	0.003	0.008	0.000
	Yes	-0.002	0.001	-0.003	-0.001	0.000	-0.000
Race/Ethnicity							
Black	No	-0.023	-0.019	-0.015	-0.028	-0.010	-0.018
	Yes	-0.020	-0.019	-0.015	-0.021	-0.011	-0.017
Hispanic	No	0.019	0.018	-0.002	-0.014	0.004	0.003
	Yes	0.013	0.003	-0.013	-0.018	-0.011	0.002
Other	No	0.013	0.018	0.005	0.003	0.002	0.010
	Yes	0.009	0.013	0.001	0.001	-0.001	0.008
Education Level							
No HS degree	No	0.002	0.002	0.000	0.003	-0.032	0.001
	Yes	-0.002	-0.004	-0.006	-0.005	-0.012	-0.002
HS degree	No	0.027	0.036	0.015	0.015	0.017	0.021
	Yes	0.012	0.023	-0.002	0.002	-0.003	0.012
Some college	No	0.001	0.002	0.005	0.007	0.012	0.015
	Yes	-0.002	0.000	0.005	0.005	0.008	0.008
College graduate	No	0.031	0.010	-0.010	-0.006	-0.004	0.015
	Yes	0.015	-0.001	-0.010	-0.008	-0.007	0.005
Age							
16 - 24 years	No	-0.011	0.034	0.010	0.012	-0.019	0.001
	Yes	-0.015	0.001	-0.004	-0.005	-0.019	0.000
25 - 44 years	No	0.000	0.003	0.001	0.001	-0.002	-0.004
	Yes	-0.001	0.001	-0.000	-0.001	-0.004	-0.004
45 - 64 years	No	0.005	-0.001	0.002	-0.003	-0.001	-0.000
	Yes	0.003	-0.004	-0.002	-0.006	-0.005	-0.000
65 years and older	No	0.023	0.028	0.007	-0.011	0.004	0.014
	Yes	0.011	0.018	-0.003	-0.014	-0.010	0.004
Country of Birth							
Not USA	No	-0.017	-0.002	-0.011	-0.002	-0.036	-0.023
	Yes	-0.019	-0.003	-0.017	-0.007	-0.030	-0.024
USA	No	0.003	0.010	0.002	0.002	0.001	0.003
	Yes	0.001	0.006	-0.001	-0.000	-0.002	0.002

Table 11-8. Document literacy intraclass correlation for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	No	0.003	0.00	0.002	0.001	0.00	0.001
	Yes	0.002	-0.001	-0.001	-0.000	0.00	0.000
Sex							
Male	No	0.006	-0.000	-0.001	0.00	0.000	0.00
	Yes	0.004	-0.002	-0.004	-0.002	-0.002	0.00
Female	No	0.003	-0.004	0.001	-0.001	-0.001	0.00
	Yes	0.002	-0.004	-0.000	-0.003	-0.003	-0.000
Census Region							
Northeast	No	0.004	-0.007	0.015	-0.003	-0.003	0.005
	Yes	0.002	0.001	0.005	-0.003	-0.003	0.004
Midwest	No	0.009	0.001	-0.002	0.009	-0.004	0.016
	Yes	0.007	-0.001	-0.002	0.004	-0.003	0.015
South	No	0.011	-0.001	0.001	0.005	0.006	0.013
	Yes	0.007	-0.001	-0.001	0.004	0.004	0.011
West	No	0.001	0.001	0.002	0.007	0.007	0.00
	Yes	0.000	-0.002	-0.000	0.001	0.001	-0.002
Race/Ethnicity							
Black	No	-0.021	-0.021	-0.010	-0.027	-0.010	-0.017
	Yes	-0.019	-0.019	-0.010	-0.025	-0.010	-0.017
Hispanic	No	0.018	-0.012	-0.001	0.007	-0.003	0.015
	Yes	0.012	-0.017	-0.010	-0.001	-0.018	0.015
Other	No	0.014	0.002	0.010	0.005	0.008	0.013
	Yes	0.009	0.001	0.002	0.002	0.001	0.011
Education Level							
No HS degree	No	0.011	0.017	0.007	0.009	0.000	0.008
	Yes	0.002	0.003	-0.006	0.002	-0.006	0.004
HS degree	No	0.023	-0.004	-0.001	0.017	0.009	0.023
	Yes	0.013	-0.005	-0.006	0.005	-0.005	0.018
Some college	No	-0.002	-0.004	0.003	0.004	0.003	0.015
	Yes	-0.002	-0.004	-0.003	-0.001	-0.001	0.014
College graduate	No	-0.006	0.010	-0.016	0.016	-0.004	0.011
	Yes	-0.007	0.001	-0.014	0.004	-0.008	0.007
Age							
16 - 24 years	No	-0.017	0.013	0.002	0.001	0.004	-0.003
	Yes	-0.018	-0.009	-0.017	-0.016	-0.010	-0.005
25 - 44 years	No	0.005	-0.002	0.002	0.003	-0.001	0.000
	Yes	0.004	-0.003	-0.002	-0.000	-0.004	-0.000
45 - 64 years	No	0.014	-0.002	-0.002	0.006	0.016	-0.002
	Yes	0.004	-0.006	-0.007	-0.000	-0.004	-0.004
65 years and older	No	0.023	0.018	0.006	-0.024	-0.011	0.019
	Yes	0.020	0.015	-0.002	-0.019	-0.014	0.010
Country of Birth							
Not USA	No	-0.017	-0.036	-0.028	0.001	-0.006	-0.020
	Yes	-0.020	-0.032	-0.028	-0.011	-0.020	-0.020
USA	No	0.005	0.001	0.003	0.002	0.004	0.004
	Yes	0.00	0.00	0.00	0.00	0.00	0.00

Table 11-9. Quantitative literacy intraclass correlation for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Imputation variance incorporated?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	No	0.00	0.00	0.00	0.00	0.00	0.00
	Yes	-0.000	-0.001	0.001	0.00	0.00	0.00
Sex							
Male	No	0.005	-0.000	0.001	0.004	0.004	-0.002
	Yes	0.002	-0.002	-0.002	-0.001	0.00	-0.002
Female	No	-0.003	0.002	0.005	0.000	-0.004	-0.002
	Yes	-0.003	-0.002	-0.001	-0.002	-0.004	-0.003
Census Region							
Northeast	No	0.001	-0.002	-0.001	-0.002	0.002	0.000
	Yes	-0.001	-0.003	-0.001	-0.002	-0.000	-0.001
Midwest	No	0.014	0.007	-0.000	0.009	-0.004	0.012
	Yes	0.008	0.001	-0.002	0.003	-0.003	0.01
South	No	0.004	0.001	0.006	0.004	-0.001	0.009
	Yes	0.003	-0.001	0.003	0.001	-0.001	0.005
West	No	-0.002	-0.001	0.002	0.010	0.000	0.001
Education Level		-0.002	-0.001	0.001	0.006	-0.000	0.001
Race/Ethnicity							
Black	No	-0.012	-0.024	-0.017	-0.023	-0.020	-0.013
	Yes	-0.012	-0.020	-0.016	-0.021	-0.019	-0.013
Hispanic	No	-0.016	-0.008	-0.005	-0.001	-0.037	0.003
	Yes	-0.016	-0.013	-0.013	-0.014	-0.031	0.000
Other	No	0.015	0.005	0.012	0.007	0.000	0.011
	Yes	0.007	0.003	0.010	0.003	-0.001	0.007
No HS degree	No	0.013	0.005	-0.010	-0.005	0.010	0.009
	Yes	0.003	-0.006	-0.010	-0.008	0.001	0.007
HS degree	No	0.017	0.019	0.012	0.017	0.004	0.012
	Yes	0.010	0.007	0.009	0.000	-0.001	0.010
Some college	No	0.002	0.000	0.003	0.014	-0.007	0.016
	Yes	0.001	-0.002	-0.002	0.001	-0.008	0.011
College graduate	No	0.007	-0.009	0.012	0.017	0.006	0.008
	Yes	0.004	-0.009	0.006	0.014	0.003	0.001
Age							
16 - 24 years	No	-0.004	-0.003	0.003	0.010	0.000	0.001
	Yes	-0.008	-0.008	-0.001	0.008	-0.003	-0.002
25 - 44 years	No	0.001	0.002	0.006	0.002	-0.001	-0.003
	Yes	-0.000	0.000	0.002	-0.001	-0.003	-0.003
45 - 64 years	No	0.014	-0.003	0.008	0.020	0.001	0.003
	Yes	0.005	-0.008	-0.002	0.008	-0.000	0.001
65 years and older	No	0.037	-0.002	0.025	-0.013	0.022	0.03
	Yes	0.025	-0.006	0.013	-0.014	0.010	0.02
Country of Birth							
Not USA	No	-0.036	-0.041	-0.027	-0.002	-0.011	-0.02
	Yes	-0.033	-0.036	-0.027	-0.013	-0.022	-0.016
USA	No	0.004	0.001	0.004	0.003	-0.001	0.00
	Yes	0.00	0.00	0.00	0.00	0.00	0.00

Table 11-10. Increase in variance attributable to variability in weights [1 + Relvar(weights)], by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	1 + Relvar (weights)
Total Population	
Total	1.34
Sex	
Male	1.30
Female	1.35
Census Region	
Northeast	1.32
Midwest	1.29
South	1.34
West	1.36
Race/Ethnicity	
Black	1.40
Hispanic	1.45
Other	1.16
Education Level	
No HS degree	1.42
HS degree	1.35
Some college	1.29
College graduate	1.21
Age	
16 - 24 years	1.44
25 - 44 years	1.32
45 - 64 years	1.30
65 years and older	1.30
Country of Birth	
Not USA	1.43
USA	1.33

11.3 COMPOSITE ESTIMATION

The basic theoretical foundation for the composite estimation indicates that when two separate samples are available for the same domain, producing two separate unbiased estimates x_1 and x_2, then $\beta x_1 + (1-\beta)x_2$ for any value of β satisfying $0 \leq \beta \leq 1$ is also an unbiased estimate. Furthermore, the variance will be minimized when

$$\beta = \frac{\sigma^2_{x_1}}{\sigma^2_{x_1} + \sigma^2_{x_2}} \qquad (11)$$

where

β = the compositing factor;

$\sigma^2_{x_1}$ = the sampling variance of x_1; and

$\sigma^2_{x_2}$ = the sampling variance of x_2.

For the National Adult Literacy Survey, x_1 is any estimate coming from the state sample for each of the eleven states involved in the survey, and x_2 is the estimate from the national part of the sample in the same state. Because the values of β for the eleven states had to be determined before any analysis of survey data could be made and values of $\sigma^2_{x_1}$ and $\sigma^2_{x_2}$ were, therefore, not available, $\sigma^2_{x_1}$ and $\sigma^2_{x_2}$ were estimated using data available from similar surveys.

Chapter 3 provides a description of the composite estimation procedure. It was expected that the resulting values of β would be reasonably close to those that minimize the sampling errors, but conclusive evidence was not available. Note that using approximations does not introduce any bias, because any value of β will produce unbiased estimates. However, if the approximations are not close to the optimum value of β, the sampling variances will be larger than is possible with a better choice of β.

The analysis that follows determines the best value of β for a variety of statistics using 1992 the National Adult Literacy Survey proficiency scores and components of variance estimated for the national and state sample designs. These variance estimates can then be used as the basis for the derivation of β for the next national assessment of adult literacy. Section 11.3.1 provides estimated compositing factors for mean proficiency and literacy levels for the five demographic variables presented in Section 11.1. Section 11.3.3 includes a model that can be used to estimate compositing factors for future national assessments of adult literacy.

11.3.1 Estimating Compositing Factors Using the National Adult Literacy Survey Data

Two sets of data files were created for the compositing analysis. One data set included the national sample cases in the PSUs within the eleven states. The second file combined data from the eleven state samples. The eleven-state national and state sample data sets were separately weighted up to the known total population following the same weighting procedures used for the National Adult Literacy Survey file. For each of the data sets, two sets of replicates were formed to compute the total and within-PSU variances. The replication designs for the total and within-PSU variances followed the guidelines given in Section 11.1.1.

The national replicates were designed to reflect the effect of the national PSU design on the estimates of variances. As mentioned earlier, the objective of the national sample design was to provide reliable statistics for the target population in the United States and for some prespecified domains of interest (i.e., regional and race/ethnicity estimates). As a result, the PSU design was not focused on producing state-level estimates, and some national sampling strata and PSUs crossed state boundaries. The replicates were designed to take into account the fact that the national sample PSUs in the 11 states were established without regard for state boundaries and that the data from these PSUs were thus subject to a higher variation than the state data.

Compositing factors were calculated for each of the 11 states as a function of the between- and within-PSU unit variances, counts of PSUs (excluding those selected for the national sample with certainty), and respondent sample sizes. The 11 state samples were combined to ensure adequate degrees of freedom for the estimation of between-PSU variances. Compositing factors were calculated separately for national certainty PSUs and the remainder of the PSUs in the sample. As mentioned earlier, because national sampling strata and PSUs crossed state boundaries, sample weights that simply reflected the reciprocal of the probabilities of selection did not provide efficient state estimates. However, this problem affected only the estimates from noncertainty PSUs.

Because the certainty PSUs in the national sample represented only themselves (i.e., a certainty PSU constituted the entire stratum), sample cases coming from these PSUs could be directly combined with the state data. Given the difference in the reliability of estimates coming from certainty and noncertainty PSUs, separate compositing factors were computed for the two types of PSUs. Separate factors were also developed for the population subgroups for which different sampling rates were used in the national sample (i.e., black, Hispanic, and other). The basic form of the composite estimator was:

$$\hat{Y}_{ik} = \beta_{ik}\hat{Y}_{(st)ik} + (1 - \beta_{ik})\hat{Y}_{(nt)ik} \tag{12}$$

where

$\hat{Y}_{ik}$ = the composite estimator for variable Y in state i for subgroup k;

β_{ik} = the compositing factor in state i for subgroup k;

$\hat{Y}_{(st)ik}$ = the estimate of Y coming from state i for subgroup k; and

$\hat{Y}_{(nt)ik}$ = the estimate of Y coming from the national sample in state i for subgroup k.

For statistic $\hat{Y}_{ik}$, the optimal compositing factor for state i and subgroup k, as given in equation (11) in Section 11.3, is

$$\beta_{ik} = \frac{V(\hat{Y}_{(nt)ik})}{V(\hat{Y}_{(nt)ik}) + V(\hat{Y}_{(st)ik})} \qquad (13)$$

where

$V(\hat{Y}_{(nt)ik})$ = the variance of the estimate of Y coming from the national sample in state i for subgroup k; and

$V(\hat{Y}_{(st)ik})$ = the variance of the estimate of Y coming from the state sample in state i for subgroup k.

For data collected in PSUs other than those selected with certainty for the national sample,

$$V(\hat{Y}_{(nt)ik}) = \frac{m_{(nc)(nt)}\,\sigma^2_{(nt)bk}}{m_{(nc)(nt)i}} + \frac{n_{(nt)k}\,\sigma^2_{(nt)wk}}{n_{(nc)(nt)ik}} \qquad (14)$$

where

$m_{(nc)(nt)}$ = the number of national sample PSUs across the eleven states that were not selected with certainty;

$\sigma^2_{(nt)bk}$ = the national between-PSU variance for subgroup k;

$m_{(nc)(nt)i}$ = the number of national sample PSUs in state i that were not selected with certainty;

$n_{(nt)k}$ = the number of respondents in the national sample across the eleven states for subgroup k;

$\sigma^2_{(nt)wk}$ = the national within-PSU variance for subgroup k; and

$n_{(nc)(nt)ik}$ = the number of national sample respondents not in national certainty PSUs in state i for subgroup k.

Similarly for states,

$$V(\hat{Y}_{(st)ik}) = \frac{m_{(nc)(st)}\,\sigma^2_{(st)bk}}{m_{(nc)(st)i}} + \frac{n_{(st)k}\,\sigma^2_{(st)wk}}{n_{(nc)(st)ik}} \qquad (15)$$

where

$m_{(nc)(st)}$ = the number of state sample PSUs across the eleven states that were not selected with certainty;

$\sigma^2_{(st)bk}$ = the state between-PSU variance for subgroup k;

$m_{(nc)(st)i}$ = the number of state sample PSUs in state i that were not selected with certainty;

$n_{(st)k}$ = the number of respondents in the state sample across the eleven states for subgroup k;

$\sigma^2_{(st)wk}$ = the state within-PSU variance for subgroup k; and

$n_{(nc)(st)ik}$ = the number of state sample respondents not in national certainty PSUs in state i for subgroup k.

For data collected in PSUs selected with certainty for the national sample, the between-PSU component of the variance is equal to 0, and the formula for variance simplifies to

$$V(\hat{Y}_{(nt)ik}) = \frac{n_{(nt)k}\,\sigma^2_{(nt)wk}}{n_{(c)(nt)ik}} \tag{16}$$

where

$n_{(c)(nt)ik}$ = the number of national sample respondents in national certainty PSUs in state i for subgroup k.

Similarly, for the state samples,

$$V(\hat{Y}_{(st)ik}) = \frac{n_{(st)k}\,\sigma^2_{(st)wk}}{n_{(c)(st)ik}} \tag{17}$$

where

$n_{(c)(st)ik}$ = the number of state sample respondents in national certainty PSUs in state i for subgroup k.

Under the assumption of equal within-PSU variance for national certainty and noncertainty PSUs, data from all PSUs were combined for the estimation of this component of variance.

Note that the specific components of variance computed in equations (14), (15), (16), and (17) reflect the national and state sample designs (i.e., the oversampling of minority populations), as well as the fact that national PSUs crossed state boundaries. States were aggregated because many states had too small national cases.

Tables 11-11, 11-12, and 11-13 provide the estimated compositing factors for average prose, document, and quantitative proficiencies and literacy levels for the state of California. The data in these tables are summarized in Figures 11-1 and 11-2 for ease of interpretation. The figures reveal the distribution of the estimated compositing factors by demographic subpopulation. Figure 11-3 shows the relatively low dispersion of the estimates of compositing factors for the state of Iowa. The individual estimated compositing factors for Iowa are shown in Tables 11-14, 11-15, and 11-16. Note that Iowa did not have any certainty PSUs in the national sample. Appendix O provides the estimated compositing factors and the associated figures for the remainder of the states in the state samples.

Table 11-11. Prose literacy optimum compositing factor for means and for percentages in each of five levels, both with and without incorporating imputation variance, by sex, census region, race/ethnicity, education level, and country of birth: California adults in 1992

		State compositing factor (Beta), including both sampling and imputation variance					
Demographic subpopulations	National certainty PSU?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	Yes	0.2700	0.3542	0.1695	0.2901	0.5171	0.2087
	No	0.3651	0.3734	0.2392	0.3308	0.5817	0.2966
Sex							
Male	Yes	0.3053	0.3134	0.2328	0.3521	0.5689	0.2387
	No	0.3389	0.3341	0.2702	0.3765	0.6337	0.2857
Female	Yes	0.3599	0.2387	0.3256	0.4144	0.3022	0.2642
	No	0.4444	0.2336	0.4179	0.4974	0.3544	0.3113
Race/Ethnicity							
Black	Yes	0.1758	0.1656	0.1948	0.0936	0.3683	0.0760
	No	0.2045	0.1876	0.2374	0.1056	0.4070	0.0968
Hispanic	Yes	0.1960	0.2921	0.2744	0.5604	0.2753	0.1776
	No	0.2298	0.2662	0.3497	0.5404	0.2478	0.1872
Other	Yes	0.4734	0.3971	0.2594	0.3358	0.6051	0.3458
	No	0.5345	0.4072	0.3057	0.3657	0.6402	0.4080
Education Level							
No HS degree	Yes	0.3163	0.4314	0.3746	0.2238	0.2561	0.1669
	No	0.3711	0.5201	0.4989	0.2905	0.3212	0.2257
HS degree	Yes	0.2863	0.2155	0.2959	0.3347	0.4641	0.2324
	No	0.4518	0.3178	0.4061	0.4734	0.6089	0.3986
Some college	Yes	0.2505	0.3212	0.1646	0.3226	0.5557	0.2862
	No	0.3041	0.3814	0.2672	0.3524	0.5878	0.3434
College graduate	Yes	0.2894	0.4628	0.4024	0.2792	0.4862	0.3973
	No	0.3227	0.5462	0.4368	0.3181	0.5556	0.4483
Age							
16 - 24 years	Yes	0.2011	0.1825	0.2447	0.4839	0.2230	0.2026
	No	0.2319	0.1965	0.3080	0.5258	0.2993	0.2916
25 - 44 years	Yes	0.2381	0.4498	0.2325	0.2981	0.5729	0.2411
	No	0.3016	0.4502	0.2837	0.3132	0.6196	0.2737
45 - 64 years	Yes	0.2684	0.3324	0.3390	0.3481	0.2939	0.2645
	No	0.3843	0.3950	0.3845	0.4192	0.3554	0.3654
Country of Birth							
Not USA	Yes	0.1557	0.2626	0.3345	0.5888	0.1894	0.1315
	No	0.1192	0.1712	0.2825	0.4653	0.1500	0.1099
USA	Yes	0.3315	0.4413	0.2080	0.2882	0.5113	0.2840
	No	0.4584	0.4949	0.2929	0.3513	0.5924	0.3495

Table 11-12. Document literacy optimum compositing factor for means and for percentages in each of five levels, within both certainty and noncertainty PSUs, by sex, census region, race/ethnicity, education level, and country of birth: California adults in 1992

Demographic subpopulations	National certainty PSU?	State compositing factor (Beta), including both sampling and imputation variance					
		Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	Yes	0.1874	0.4016	0.4160	0.7227	0.6330	0.2622
	No	0.2415	0.4981	0.4885	0.7197	0.6767	0.3270
Sex							
Male	Yes	0.2925	0.3692	0.4357	0.5673	0.4119	0.3087
	No	0.2833	0.3975	0.4491	0.5604	0.4451	0.3341
Female	Yes	0.1763	0.4063	0.2193	0.4315	0.4694	0.2509
	No	0.2687	0.4921	0.3188	0.5017	0.5274	0.2879
Race/Ethnicity							
Black	Yes	0.1354	0.2122	0.2029	0.1489	0.0984	0.0838
	No	0.1591	0.2392	0.2437	0.1680	0.1088	0.1049
Hispanic	Yes	0.1811	0.2740	0.2695	0.3522	0.4480	0.1528
	No	0.1644	0.2579	0.2471	0.3338	0.3917	0.1649
Other	Yes	0.3393	0.4394	0.4924	0.6799	0.6839	0.4254
	No	0.4075	0.5295	0.5543	0.6685	0.7163	0.4709
Education Level							
No HS degree	Yes	0.2391	0.3105	0.3557	0.2613	0.1895	0.1944
	No	0.2846	0.3544	0.4199	0.3349	0.2425	0.2530
HS degree	Yes	0.2056	0.2584	0.3303	0.3499	0.3451	0.2447
	No	0.3338	0.4015	0.4662	0.4981	0.5124	0.4120
Some college	Yes	0.1754	0.5049	0.6857	0.4286	0.5189	0.2953
	No	0.1908	0.5363	0.6764	0.4537	0.4729	0.3129
College graduate	Yes	0.3383	0.4349	0.2402	0.6086	0.5490	0.4355
	No	0.3840	0.4672	0.3167	0.6730	0.6386	0.5172
Age							
16 – 24 years	Yes	0.2489	0.5922	0.4916	0.2979	0.3271	0.1977
	No	0.2852	0.5698	0.5592	0.3354	0.3808	0.2646
25 – 44 years	Yes	0.1226	0.2466	0.3090	0.6763	0.6211	0.2964
	No	0.1836	0.3413	0.3882	0.6844	0.6732	0.3243
45 – 64 years	Yes	0.3726	0.3259	0.3331	0.4111	0.5115	0.3583
	No	0.4445	0.4254	0.4309	0.5055	0.5947	0.4457
Country of Birth							
Not USA	Yes	0.2043	0.2396	0.3275	0.4768	0.3815	0.1760
	No	0.1381	0.1414	0.2348	0.4092	0.3224	0.1440
USA	Yes	0.1797	0.3390	0.4056	0.6623	0.6333	0.3449
	No	0.2548	0.4633	0.4944	0.6804	0.6906	0.3900

Table 11-13. Quantitative literacy optimum compositing factor for means and for percentages in each of five levels, within both certainty and noncertainty PSUs, by sex, census, race/ethnicity, education level, and country of birth: California adults in 1992

		State compositing factor (Beta), including both sampling and imputation variance					
Demographic subpopulations	National certainty PSU?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	Yes	0.1197	0.2967	0.3626	0.4277	0.4839	0.1935
	No	0.2047	0.3528	0.3770	0.4977	0.5548	0.2617
Sex							
Male	Yes	0.2642	0.2398	0.4776	0.4690	0.4473	0.2130
	No	0.2966	0.2477	0.4783	0.5063	0.5160	0.2245
Female	Yes	0.1970	0.5792	0.3958	0.3847	0.3541	0.2479
	No	0.2902	0.6287	0.4616	0.4721	0.3859	0.2964
Race/Ethnicity							
Black	Yes	0.1665	0.1862	0.1522	0.1405	0.3526	0.0729
	No	0.2086	0.2158	0.1744	0.1609	0.3908	0.0966
Hispanic	Yes	0.1493	0.3891	0.2386	0.3555	0.0650	0.1567
	No	0.1638	0.3740	0.2571	0.4159	0.0338	0.1900
Other	Yes	0.3384	0.3675	0.4087	0.4518	0.5617	0.4265
	No	0.4388	0.4217	0.4168	0.5093	0.5977	0.4517
Education Level							
No HS degree	Yes	0.1184	0.2926	0.1676	0.2119	0.1851	0.1305
	No	0.1901	0.2798	0.2435	0.3082	0.2139	0.2152
HS degree	Yes	0.2028	0.2685	0.2377	0.2374	0.6514	0.1725
	No	0.3669	0.4416	0.3420	0.3607	0.7237	0.3240
Some college	Yes	0.2248	0.1894	0.4217	0.6094	0.4189	0.2841
	No	0.2713	0.2313	0.4453	0.6370	0.4633	0.2883
College graduate	Yes	0.2096	0.2935	0.3398	0.2227	0.4271	0.5283
	No	0.2459	0.3277	0.3652	0.2961	0.4749	0.5595
Age							
16 – 24 years	Yes	0.2033	0.2934	0.1924	0.1472	0.2527	0.2206
	No	0.2803	0.3344	0.2395	0.1857	0.3397	0.3219
25 – 44 years	Yes	0.1431	0.3756	0.4048	0.5832	0.5684	0.1923
	No	0.2318	0.4302	0.4396	0.6157	0.6133	0.2315
45 – 64 years	Yes	0.3086	0.2120	0.3219	0.4763	0.3645	0.3210
	No	0.3900	0.2737	0.3695	0.5781	0.4009	0.4164
Country of Birth							
Not USA	Yes	0.1672	0.1678	0.2515	0.5612	0.3532	0.1730
	No	0.1229	0.1156	0.1622	0.4573	0.2499	0.1334
USA	Yes	0.1939	0.3267	0.4379	0.4558	0.5210	0.2470
	No	0.2839	0.4092	0.4878	0.5560	0.6100	0.3302

Table 11-14. Prose literacy optimum compositing factor for means and for percentages in each of five levels, within both certainty and noncertainty PSUs, by sex, census region, race/ethnicity, education level, and country of birth: Iowa adults in 1992

		State compositing factor (Beta), including both sampling and imputation variance					
Demographic subpopulations	National certainty PSU?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	Yes	-	-	-	-	-	-
	No	0.8301	0.8405	0.7289	0.8185	0.9259	0.7849
Sex							
Male	Yes	-	-	-	-	-	-
	No	0.8069	0.8145	0.7539	0.8362	0.9367	0.7617
Female	Yes	-	-	-	-	-	-
	No	0.8753	0.7310	0.8676	0.9036	0.8351	0.7958
Race/Ethnicity							
Black	Yes	-	-	-	-	-	-
	No	0.9481	0.9439	0.9537	0.8979	0.9803	0.8745
Hispanic	Yes	-	-	-	-	-	-
	No	0.8717	0.9207	0.9148	0.9729	0.9128	0.8542
Other	Yes	-	-	-	-	-	-
	No	0.8607	0.7694	0.7015	0.7449	0.9010	0.7891
Education Level							
No HS degree	Yes	-	-	-	-	-	-
	No	0.9056	0.9422	0.9347	0.8628	0.8820	0.8172
HS degree	Yes	-	-	-	-	-	-
	No	0.8903	0.8121	0.8636	0.8939	0.9375	0.8685
Some college	Yes	-	-	-	-	-	-
	No	0.8239	0.8600	0.7743	0.8555	0.9332	0.8395
College graduate	Yes	-	-	-	-	-	-
	No	0.6161	0.8106	0.7233	0.6050	0.8044	0.7303
Age							
16 – 24 years	Yes	-	-	-	-	-	-
	No	0.6969	0.6420	0.7677	0.8917	0.7588	0.7534
25 – 44 years	Yes	-	-	-	-	-	-
	No	0.8000	0.8932	0.7908	0.8257	0.9412	0.7820
45 – 64 years	Yes	-	-	-	-	-	-
	No	0.8402	0.8423	0.8386	0.8589	0.8203	0.8304
Country of Birth							
Not USA	Yes	-	-	-	-	-	-
	No	0.5431	0.6902	0.7658	0.8999	0.6024	0.4971
USA	Yes	-	-	-	-	-	-
	No	0.8637	0.8752	0.7559	0.8001	0.9145	0.7957

Table 11-15. Document literacy optimum compositing factor for means and for percentages in each of five levels, within both certainty and noncertainty PSUs, by sex, census region, race/ethnicity, education level, and country of birth: Iowa adults in 1992

		State compositing factor (Beta), including both sampling and imputation variance					
Demographic subpopulations	National certainty PSU?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	Yes	-	-	-	-	-	-
	No	0.7224	0.8976	0.8934	0.9556	0.9497	0.8063
Sex							
Male	Yes	-	-	-	-	-	-
	No	0.7598	0.8434	0.8655	0.9091	0.8711	0.7995
Female	Yes	-	-	-	-	-	-
	No	0.7591	0.8941	0.8101	0.9027	0.9123	0.7794
Race/Ethnicity							
Black	Yes	-	-	-	-	-	-
	No	0.9292	0.9578	0.9566	0.9365	0.9023	0.8855
Hispanic	Yes	-	-	-	-	-	-
	No	0.8557	0.9123	0.9091	0.9379	0.9572	0.8293
Other	Yes	-	-	-	-	-	-
	No	0.7909	0.8648	0.8700	0.9062	0.9281	0.8236
Education Level							
No HS degree	Yes	-	-	-	-	-	-
	No	0.8575	0.9018	0.9207	0.8814	0.8351	0.8355
HS degree	Yes	-	-	-	-	-	-
	No	0.8270	0.8697	0.8971	0.9001	0.9167	0.8691
Some college	Yes	-	-	-	-	-	-
	No	0.7170	0.9223	0.9538	0.8924	0.9029	0.8240
College graduate	Yes	-	-	-	-	-	-
	No	0.6759	0.7555	0.6278	0.8758	0.8565	0.7949
Age							
16 – 24 years	Yes	-	-	-	-	-	-
	No	0.7498	0.9063	0.9039	0.7884	0.8211	0.7310
25 – 44 years	Yes	-	-	-	-	-	-
	No	0.6643	0.8289	0.8642	0.9548	0.9534	0.8157
45 – 64 years	Yes	-	-	-	-	-	-
	No	0.8658	0.8651	0.8636	0.8934	0.9259	0.8697
Country of Birth							
Not USA	Yes	-	-	-	-	-	-
	No	0.6145	0.6577	0.7537	0.8555	0.8003	0.5807
USA	Yes	-	-	-	-	-	-
	No	0.7078	0.8679	0.8770	0.9353	0.9421	0.8185

Table 11-16. *Quantitative literacy optimum* compositing factor for means and for percentages in each of five levels, within both certainty and noncertainty PSUs, by sex, census region, race/ethnicity, education level, and country of birth: Iowa adults in 1992

		State compositing factor (Beta), including both sampling and imputation variance					
Demographic subpopulations	National certainty PSU?	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	Yes	-	-	-	-	-	-
	No	0.6711	0.8254	0.8450	0.9005	0.9190	0.7534
Sex							
Male	Yes	-	-	-	-	-	-
	No	0.7654	0.7235	0.8849	0.8944	0.9010	0.6962
Female	Yes	-	-	-	-	-	-
	No	0.7787	0.9391	0.8893	0.8939	0.8524	0.7858
Race/Ethnicity							
Black	Yes	-	-	-	-	-	-
	No	0.9444	0.9508	0.9389	0.9329	0.9788	0.8710
Hispanic	Yes	-	-	-	-	-	-
	No	0.8288	0.9461	0.8983	0.9407	0.6598	0.8383
Other	Yes	-	-	-	-	-	-
	No	0.8202	0.7953	0.7743	0.8467	0.8836	0.8064
Education Level							
No HS degree	Yes	-	-	-	-	-	-
	No	0.7577	0.8739	0.8220	0.8713	0.8285	0.7865
HS degree	Yes	-	-	-	-	-	-
	No	0.8485	0.8890	0.8250	0.8470	0.9547	0.8291
Some college	Yes	-	-	-	-	-	-
	No	0.7946	0.7522	0.8906	0.9473	0.8968	0.8076
College graduate	Yes	-	-	-	-	-	-
	No	0.5253	0.6215	0.6340	0.6142	0.7545	0.8090
Age							
16 – 24 years	Yes	-	-	-	-	-	-
	No	0.7413	0.7835	0.6998	0.6320	0.7896	0.7780
25 – 44 years	Yes	-	-	-	-	-	-
	No	0.7268	0.8792	0.8830	0.9404	0.9395	0.7343
45 – 64 years	Yes	-	-	-	-	-	-
	No	0.8384	0.7600	0.8299	0.9203	0.8494	0.8513
Country of Birth							
Not USA	Yes	-	-	-	-	-	-
	No	0.5609	0.5602	0.6749	0.8899	0.7719	0.5731
USA	Yes	-	-	-	-	-	-
	No	0.7403	0.8334	0.8739	0.9054	0.9215	0.7866

11.3.2 Evaluating the National Adult Literacy Survey Compositing Factors

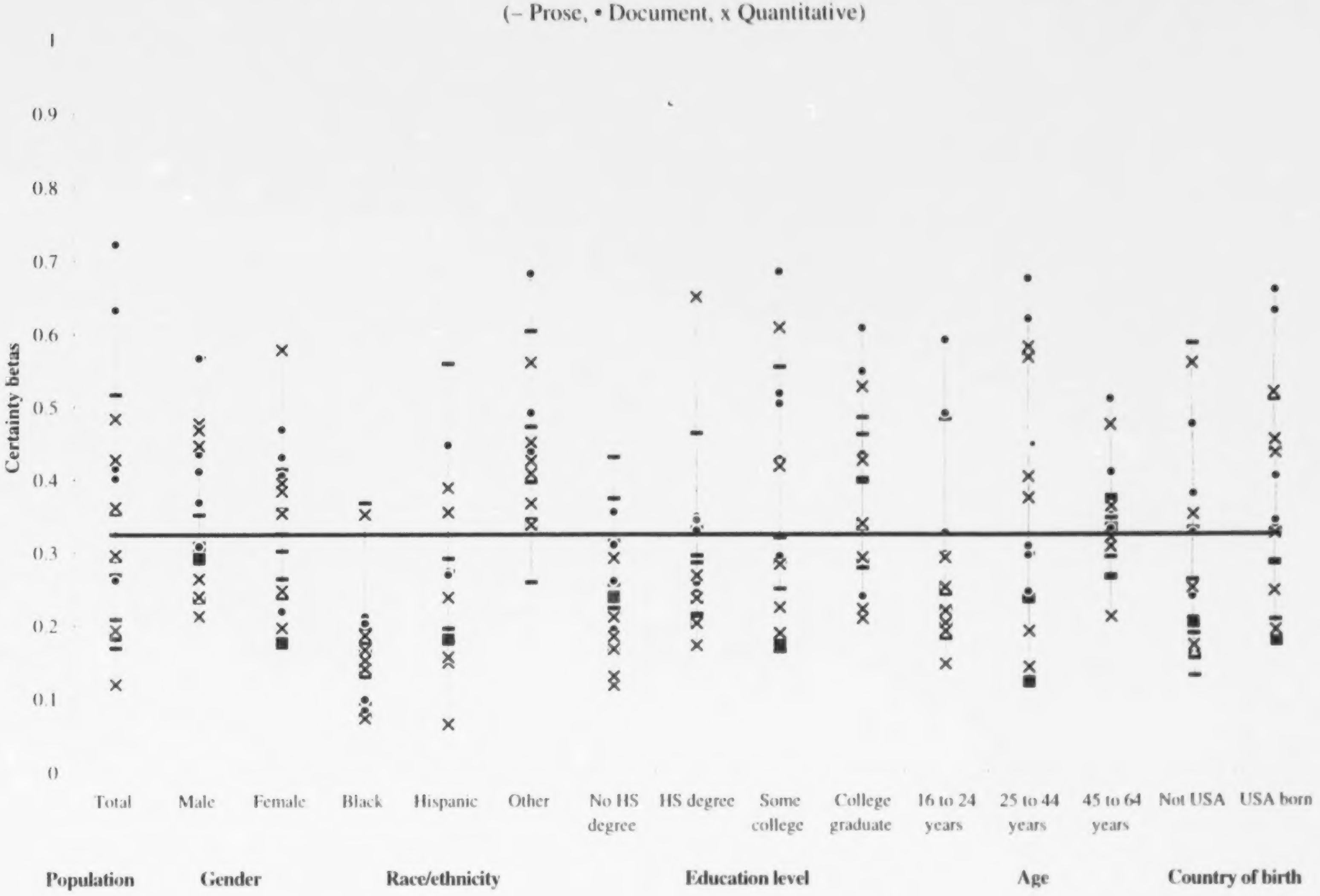
The main objective of compositing the national and state samples was to improve the precision of the estimates. The composite estimation did improve the statistics coming from the eleven state samples. It also improved the precision of statistics coming from the national sample, but the relative gain was lower than for the eleven states. Tables 11-17, 11-18, and 11-19 show the percent decrease in variance for national statistics after compositing the national and state data. The tables also present the percent increase in the sample size after compositing the data. The general pattern indicates that the variances were decreased as a result of compositing but at a much lower rate than the increase in sample sizes. This is not a surprising outcome because the additional sample size came from eleven states that made up about one-half of the total U.S. population. In some cases, the percent decrease is a negative number, indicating that variances were increased as a result of compositing. It should be noted, however, that the variances of some of the items in the tables are quite small, making the ratio (the estimate of the percent decrease in variance) very unstable. For example, in Table 11-17 the estimated variances of level 3 prose literacy scores for males are 0.000105 and 0.000145 before and after compositing, respectively. The difference between the two estimates is trivial, even though the table shows a 38 percent increase in the variance. The same is true in table 11-19 for level 2 quantitative scores for persons born outside of the United States; the estimated variances before and after compositing are 0.0001716 and 0.0002683, respectively.

Another factor that should be considered when studying these tables is that the entries are estimates themselves and are subject to variation. The variances for overall document proficiency for the South are 3.358 and 4.174 before and after compositing, respectively resulting in a decrease of –17% in the variance shown in Table 11-18. These figures are estimates with some associated variability.

11.3.3 Compositing Factors for a Future Assessment of Adult Literacy

An objective of the compositing analysis was to use the National Adult Literacy Survey data to arrive at a method of estimating the compositing factors for the next cycle of adult literacy assessment. As was shown in the previous section, estimates of optimum compositing factors depend on the state and national sample designs and sample sizes, as well as on the specific statistics under study. Different optimal values of the compositing factors exist for each statistic of interest. However, data analyses would become quite complicated if item-specific values of the compositing factors were used. Subgroup estimates would not necessarily add to totals, statistics obtained in different ways would be inconsistent, and it would not be clear what the factors should be for items for which compositing factors had not been calculated. Consequently, as in the first cycle of the National Adult Literacy Survey, the goal of the evaluation was to associate with each sample person a single compositing factor that, while not precisely optimal for any particular statistic, would be robust enough to enhance the precision of virtually all composited statistics.

Figure 11-1. California Certainty Betas



2G7

266

Figure 11-2. California Noncertainty Betas

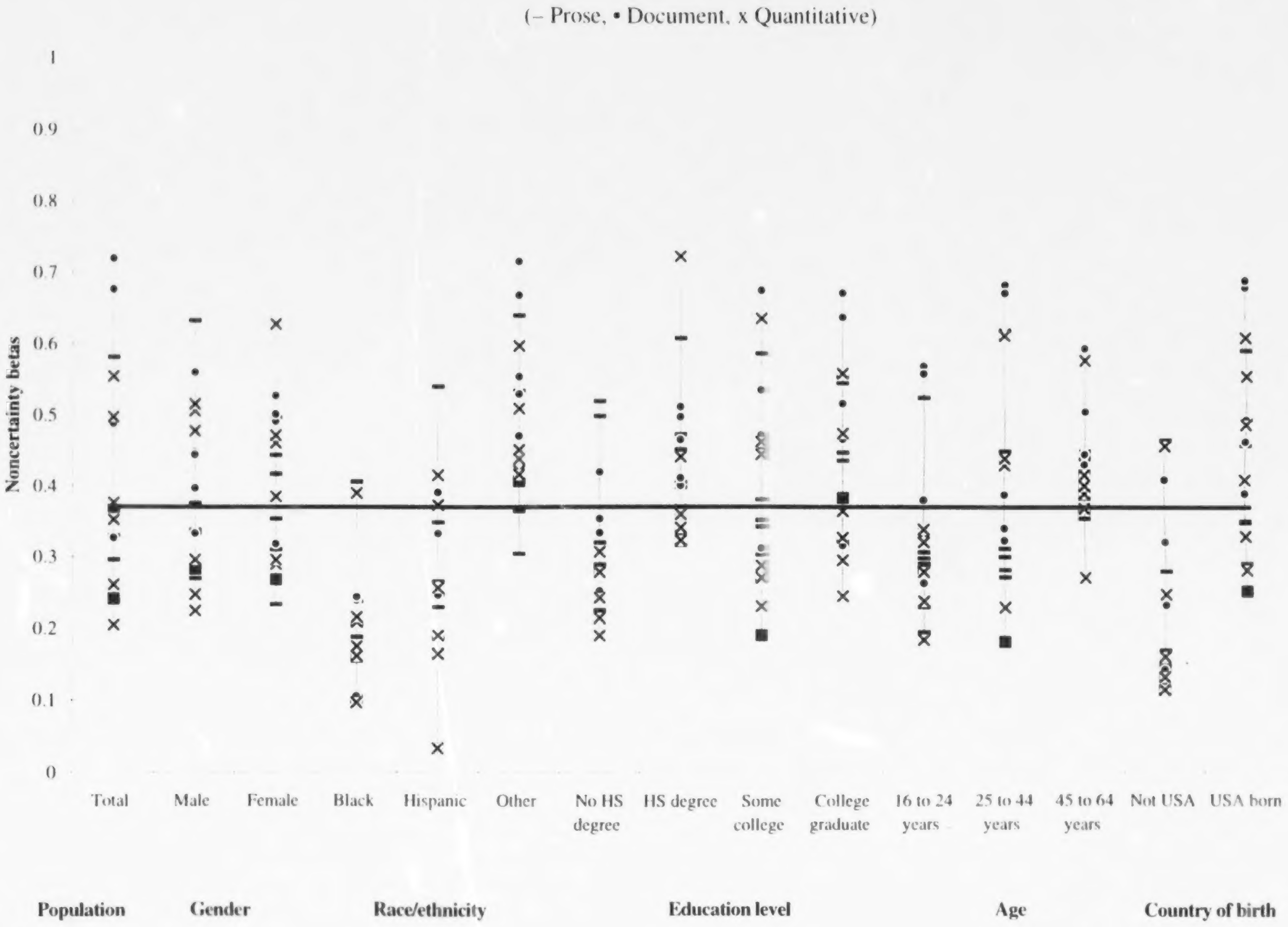
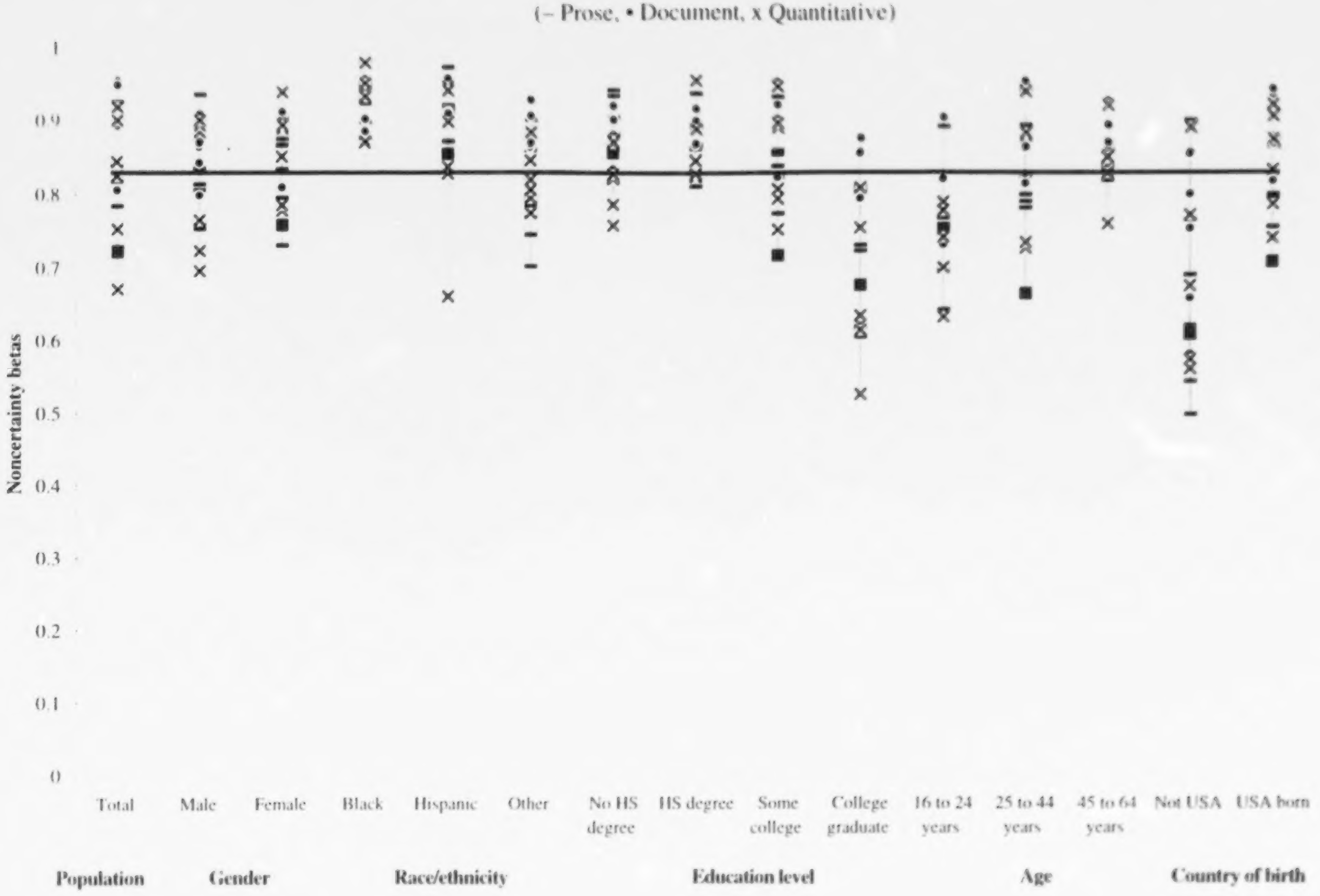


Figure 11-3. Iowa Noncertainty Betas



270

271

Table 11-17. Prose literacy percent change in sample size and variance after compositing for means and for percentages in each of five levels, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

		Percent Decrease in Variance After Compositing, including imputation variance as well as sampling variance					
Demographic subpopulations	Percent increase in sample size	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	83.5	38.5	17.0	-11.9	12.1	50.5	18.7
Sex							
Male	87.7	31.6	-3.6	-38.0	42.0	35.7	16.6
Female	80.6	38.0	9.2	52.0	15.3	40.0	20.0
Census Region							
Northeast	113.9	61.1	11.9	58.5	53.2	38.0	62.4
Midwest	137.2	39.6	32.3	16.0	19.6	57.7	44.2
South	43.6	-19.3	10.5	11.7	3.2	9.2	-26.9
West	73.0	-1.2	-2.1	-53.9	39.4	61.7	-3.7
Race/Ethnicity							
Black	41.4	16.4	-31.9	13.5	4.8	36.5	11.5
Hispanic	46.6	44.9	43.0	27.2	40.7	26.1	35.0
Other	108.1	50.9	32.5	0.1	22.8	52.3	36.7
Education Level							
No HS degree	55.9	22.4	46.3	15.8	15.7	-5.1	19.7
HS degree	88.0	43.4	22.8	4.4	-1.2	44.9	25.5
Some college	95.6	30.5	31.1	18.9	27.9	37.5	29.9
College graduate	101.7	20.0	10.9	-2.8	25.2	57.1	25.3
Age							
16 – 24 years	88.0	18.5	35.4	18.9	21.2	23.4	38.7
25 – 44 years	97.7	54.2	23.7	1.2	19.0	39.0	29.1
45 – 64 years	108.6	34.6	27.7	15.4	18.6	39.6	2.6
Country of Birth							
Not USA	63.8	23.1	36.5	11.5	33.7	26.8	6.2
USA	86.1	26.5	18.3	-9.8	17.6	49.5	6.3

Table 11-18. Document Literacy percent change in sample size and variance after compositing for means and for percentages in each of five levels, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

		Percent Decrease in Variance After Compositing, including imputation variance as well as sampling variance					
Demographic Subpopulations	Percent increase in sample size	Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average Proficiency
Total Population							
Total	83.5	22.2	31.6	47.2	30.8	50.2	12.3
Gender							
Male	87.7	19.4	45.6	45.1	45.4	30.6	16.5
Female	80.6	12.5	8.5	25.6	37.0	48.8	13.2
Census Region							
Northeast	113.9	53.5	52.8	64.5	26.9	0.2	52.0
Midwest	137.2	41.0	22.5	43.2	37.0	23.7	49.3
South	43.6	-6.2	11.2	17.2	1.4	17.9	-17.5
West	73.0	-2.5	22.8	27.7	51.9	50.4	-12.8
Race/Ethnicity							
Black	41.4	30.5	2.7	28.1	5.0	37.2	32.7
Hispanic	46.6	23.9	26.6	18.6	19.8	45.5	27.7
Other	108.1	29.6	13.2	44.1	29.1	50.5	29.2
Education Level							
No HS degree	55.9	9.4	6.7	32.9	45.5	-0.3	12.5
Country of Birth							
Not USA	63.8	28.9	30.8	12.9	40.4	72.5	13.0
USA	86.1	16.7	33.9	46.9	24.6	50.7	11.6

Table 11-19. Quantitative literacy percent change in sample size and variance after compositing for means and for percentages in each of five levels, by sex, census region, race/ethnicity, education level, and country of birth: Adults in 1992

Demographic subpopulations	Percent increase in sample size	Percent Decrease in Variance After Compositing, including imputation variance as well as sampling variance					
		Percent in Level 1 (225 or lower)	Percent in Level 2 (226 to 275)	Percent in Level 3 (276 to 325)	Percent in Level 4 (326 to 375)	Percent in Level 5 (376 or higher)	Average proficiency
Total Population							
Total	83.5	13.2	-12.6	-3.4	59.0	43.1	16.9
Sex							
Male	87.7	11.9	45.5	54.4	70.5	46.9	25.8
Female	80.6	30.6	17.2	17.8	25.4	-3.5	15.2
Census Region							
Northeast	113.9	67.5	69.8	52.2	56.2	45.3	54.6
Midwest	137.2	31.8	7.8	4.1	48.5	63.2	14.0
South	43.6	-17.2	19.7	-0.9	12.6	1.1	-24.3
West	73.0	-27.4	-9.7	40.0	41.9	23.9	-8.4
Race/Ethnicity							
Black	41.4	56.6	24.0	5.4	10.8	12.4	36.5
Hispanic	46.6	25.0	33.4	35.3	17.2	11.4	29.3
Other	108.1	31.0	5.1	14.6	58.7	42.4	32.8
Education Level							
No HS degree	55.9	-4.1	-5.2	-24.2	38.6	-1.9	11.5
HS degree	88.0	31.5	24.3	1.3	62.9	53.3	20.2
Some college	95.6	11.2	-43.6	3.5	24.3	43.8	16.8
College graduate	101.7	-10.2	47.7	34.3	25.2	24.0	41.0
Age							
16 – 24 years	88.0	20.2	17.0	30.8	9.2	-46.0	37.5
25 – 44 years	97.7	37.2	-0.7	28.8	70.8	47.6	30.4
45 – 64 years	108.6	5.9	6.2	45.8	31.0	27.9	20.0
Country of Birth							
Not USA	63.8	-6.0	-56.4	15.8	56.8	29.3	-1.2
USA	86.1	12.3	1.5	1.4	55.4	41.4	15.4

The objective can be accomplished by focusing on aspects of the sample design that are likely to affect the variance, regardless of the choice of statistic.

There are two aspects of the design that should be specifically reflected in the compositing factors. One is the distinction between cases coming from national certainty or noncertainty PSUs, as is apparent in equations (15), (16), (17), (18), and (19). The next design aspect is the oversampling of black and Hispanic respondents in the national sample. The oversampling introduced variability in the weights and increased the design effect for cases coming from the national sample. To best reflect this design feature, there should be separate compositing factors for minority and nonminority groups as in the first cycle of National Adult Literacy Survey (refer to Section 11.2.4 for details).

The compositing factor equation (14) can be rewritten in terms of the ratios of the variances coming from national and state samples as

$$\beta_{ik} = \frac{\dfrac{V(\hat{Y}_{(nt)ik})}{V(\hat{Y}_{(st)ik})}}{\dfrac{V(\hat{Y}_{(nt)ik})}{V(\hat{Y}_{(st)ik})} + 1} = \frac{\text{Ratio}_{ik}(\text{var})}{\text{Ratio}_{ik}(\text{var}) + 1}$$

where

$$\text{Ratio}_{ik}(\text{var}) = \frac{V(\hat{Y}_{(nt)ik})}{V(\hat{Y}_{(st)ik})} \tag{18}$$

For sample cases coming from the national certainty PSUs,

$$\text{Ratio}_{ik}(\text{var}) = \frac{n_{(c)(st)ik}}{n_{(c)(nt)ik}} \times R_{(c)ik} \tag{19}$$

where, following equations (15) and (16),

$$R_{(c)ik} = \frac{n_{(c)(nt)ik} \quad \sigma^2_{(nt)wk}}{n_{(c)(st)ik} \quad \sigma^2_{(st)wk}} = \frac{\sigma^2_{u(c)(nt)k}}{\sigma^2_{u(c)(st)k}}$$

$\sigma^2_{u(c)(nt)k}$ = the national unit within-PSU variance for subgroup k and;

$\sigma^2_{u(c)(st)k}$ = the state unit within-PSU variance for subgroup k.

The remainder of the terms in equation (19) are as defined in equations (16) and (17).

For sample cases not in national certainty PSUs,

$$\text{Ratio}_{ik}(\text{var}) = R_{(nc)ik} \times \left\{ \frac{\dfrac{P_{(nt)ik}}{m_{(nc)(nt)i}} + \dfrac{1 - P_{(nt)ik}}{n_{(nc)(nt)ik}}}{\dfrac{P_{(st)ik}}{m_{(nc)(st)i}} + \dfrac{1 - P_{(st)ik}}{n_{(nc)(st)ik}}} \right\} \tag{20}$$

where, following equations (14) and (15),

$$R_{(nc)ik} = \frac{m_{(nc)(nt)}\,\sigma^2_{(nt)bk} + n_{(nt)}\,\sigma^2_{(nt)wk}}{m_{(nc)(st)}\,\sigma^2_{(st)bk} + n_{(st)ik}\,\sigma^2_{(st)wk}} = \frac{\sigma^2_{u(nc)(nt)k}}{\sigma^2_{u(nc)(st)k}} \tag{21}$$

$$P_{(nt)k} = \frac{m_{(nc)(nt)}\,\sigma^2_{(nt)bk}}{m_{(nc)(nt)}\,\sigma^2_{(nt)bk} + n_{(nt)k}\,\sigma^2_{(nt)wk}}$$

$$P_{(st)k} = \frac{m_{(nc)(st)}\,\sigma^2_{(st)bk}}{m_{(nc)(nt)}\,\sigma^2_{(st)bk} + n_{(nt)k}\,\sigma^2_{(st)wk}}$$

$\sigma^2_{u(nc)(nt)k}$ = the national unit variance for cases not in national certainty PSUs for subgroup k;

$\sigma^2_{u(nc)(st)k}$ = the state unit variance for cases not in national certainty PSUs for subgroup k;

$P_{(nt)k}$ = the proportion of the national unit variance for subgroup k coming from the between-PSU component; and

$P_{(st)k}$ = the proportion of the state unit variance for subgroup k coming from the between-PSU component.

The remainder of the terms in equation (20) are as defined in equations (14) and (15).

Because a key objective of this analysis was to produce estimates of compositing factors that are constant across various statistics, the individual values of the statistics in equations (19) and (20) were replaced by their average values, as described in the following model:

$$\beta_{ik} = \frac{\overline{\text{Ratio}}_{ik}(\text{var})}{\overline{\text{Ratio}}_{ik}(\text{var}) + 1} \tag{22}$$

where

$\overline{\beta}_{ik}$ = an estimate of the compositing factor for state i and subgroup k.

For sample cases in the national certainty PSUs,

$$\overline{\text{Ratio}}_{i(other)}(\text{var}) = \frac{n_{(c)(st)i(other)}}{n_{(c)(nt)(other)}} \times \bar{R}_{(c)ik} \tag{23}$$

where

$\overline{R}_{(c)ik}$ = the average value of the ratio of the unit variances for sample cases in national certainty PSUs.

For sample cases not in national certainty PSUs,

$$\overline{\text{Ratio}}_{ik}(\text{var}) = \bar{R}_{(nc)ik} \times \left\{ \frac{\dfrac{\bar{P}_{(nt)ik}}{m_{(nc)(nt)i}} + \dfrac{1-\bar{P}_{(nt)ik}}{n_{(nc)(nt)ik}}}{\dfrac{\bar{P}_{(st)ik}}{m_{(nc)(st)i}} + \dfrac{1-\bar{P}_{(st)ik}}{n_{(nc)(st)ik}}} \right\} \tag{24}$$

where

$\overline{R}_{(nc)ik}$ = the average value of the ratio of the unit variances for sample cases in PSUs that were not national certainty PSUs;

$\overline{P}_{(nt)k}$ = the average value of $P_{(nt)k}$ as defined in equation (21); and

$\overline{P}_{(st)k}$ = the average value of $P_{(st)k}$ as defined in equation (21).

The average values of the ratio of the unit variances in equations (20) and (21) were computed separately for black, Hispanic, and other race/ethnicity groups for mean proficiency and the five literacy levels (as defined in Section 11.1) for prose, document, and quantitative literacy scores. The values of $P_{(nt)k}$ and $P_{(st)k}$ were also averaged over the same statistics. Substituting the values of $\overline{R}_{(c)ik}$, $\overline{R}_{(nc)ik}$, $\overline{P}_{(nt)k}$, and $\overline{P}_{(st)k}$ in equations (23) and (24) will give the following models:

$$\overline{\text{Ratio}}_{i(\text{Black})}(\text{var}) = \frac{n_{(c)(st)i(\text{Black})}}{n_{(c)(nt)i(\text{Black})}} \times 0.7363$$

$$\overline{\text{Ratio}}_{i(\text{Hispanic})}(\text{var}) = \frac{n_{(c)(st)i(\text{Hispanic})}}{n_{(c)(nt)i(\text{Hispanic})}} \times 1.1275 \tag{25}$$

$$\overline{\text{Ratio}}_{i(\text{Other})}(\text{var}) = \frac{n_{(c)(st)i(\text{Other})}}{n_{(c)(nt)i(\text{Other})}} \times 0.8118$$

The estimated compositing factor for cases in national certainty PSUs can be computed by replacing the values of $\overline{\text{Ratio}}_{(c)ik}(\text{var})$ as computed in equation (22).

$$\overline{\text{Ratio}}_{(nc)i(\text{Black})}(\text{var}) = 0.7365 \times \left\{ \frac{\dfrac{0.0054}{m_{(nc)(nt)i}} + \dfrac{0.9946}{n_{(nc)(nt)i(\text{Black})}}}{\dfrac{0.0041}{m_{(nc)(st)i}} + \dfrac{0.9959}{n_{(nc)(st)i(\text{Black})}}} \right\}$$

$$\overline{\text{Ratio}}_{(nc)i(\text{Hispanic})}(\text{var}) = 1.1322 \times \left\{ \frac{\dfrac{0.0032}{m_{(nc)(nt)i}} + \dfrac{0.9968}{n_{(nc)(nt)i(\text{Hispanic})}}}{\dfrac{0.0004}{m_{(nc)(st)i}} + \dfrac{0.9996}{n_{(nc)(st)i(\text{Hispanic})}}} \right\} \qquad (26)$$

$$\overline{\text{Ratio}}_{(nc)i(\text{Other})}(\text{var}) = 0.8115 \times \left\{ \frac{\dfrac{0.0008}{m_{(nc)(nt)i}} + \dfrac{0.9992}{n_{(nc)(nt)i(\text{Other})}}}{\dfrac{0.0006}{m_{(nc)(st)i}} + \dfrac{0.9994}{n_{(nc)(st)i(\text{Other})}}} \right\}$$

The estimated compositing factor for cases not in national certainty PSUs can be computed by replacing the values of $\overline{\text{Ratio}}_{(nc)ik}(\text{var})$ in equation (24).

Table 11-20 examines the performance of the compositing factors estimated above by comparing them to the factors precomputed for the 1992 National Adult Literacy Survey data. In general, there is a very strong agreement between the two sets of compositing factors. The only noticeable differences occur in cases where the sample sizes were too small to produce a reliable estimate (e.g., Hispanic respondents not in national certainty PSUs in New York).

Because the compositing factors were based on the sample designs used in the 1992 national and state samples, they will have to be modified if different designs are applied in future national assessments of adult literacy. For example, different compositing factors were used for minority and nonminority groups mainly because minority groups were oversampled in the national component only. If the next design does not involve oversampling minority groups, then the compositing factors computed for the non-black, non-Hispanic race/ethnicity group can be used for both minority and nonminority groups. If the next design uses different numbers of PSUs or different average segment sizes, oversamples other

Table 11-20. Comparison of compositing factors resulting from model to those used in the National Adult Literacy Survey

State	Certainty PSU?	Race/Ethnicity	Data source	Noncertainty PSU count	Respondent count	Compositing Factor: From model	Compositing Factor: Used in survey
California	No	Black	National	4	26	0.7929	0.7098
			State	4	9	0.2071	0.2902
		Hispanic	National	4	182	0.7093	0.7098
			State	4	58	0.2907	0.2902
		Other	National	4	188	0.4760	0.4401
			State	4	255	0.5240	0.5599
	Yes	black	National	-	219	0.8229	0.6883
			State	-	64	0.1771	0.3117
		Hispanic	National	-	470	0.7103	0.6883
			State	-	170	0.2897	0.3117
		Other	National	-	400	0.5232	0.5232
			State	-	449	0.4768	0.4768
Illinois	No	Black	National	3	10	0.4043	0.5968
			State	5	20	0.5957	0.4032
		Hispanic	National	3	49	0.8049	0.5968
			State	5	10	0.1951	0.4032
		Other	National	3	199	0.3699	0.3210
			State	5	416	0.6301	0.6790
	Yes	Black	National	-	102	0.5297	0.4378
			State	-	123	0.4703	0.5622
		Hispanic	National	-	60	0.4150	0.4378
			State	-	75	0.5850	0.5622
		Other	National	-	120	0.2811	0.2502
			State	-	378	0.7189	0.7498
Indiana	No	Black	National	4	99	0.5309	0.3834
			State	9	110	0.4691	0.6166
		Hispanic	National	4	8	0.2930	0.3834
			State	9	17	0.7070	0.6166
		Other	National	4	215	0.2219	0.1746
			State	9	946	0.7781	0.8254
Iowa	No	Black	National	2	1	0.0519	0.0441
			State	11	25	0.9481	0.9559
		Hispanic	National	2	2	0.0812	0.0441
			State	11	20	0.9188	0.9559
		Other	National	2	145	0.1478	0.1073
			State	11	1027	0.8522	0.8927
Louisiana	No	Black	National	1	76	0.2313	0.1559
			State	7	283	0.7687	0.8441

Table 11-20. Comparison of compositing factors resulting from model to those used in the National Adult Literacy Survey — Continued

		Hispanic	National	1	4	0.0887	0.1559
			State	7	36	0.9113	0.8441
		Other	National	1	55	0.0879	0.0649
			State	7	714	0.9121	0.9351
New Jersey	No	Black	National	4	80	0.4668	0.3293
			State	-	112	0.5332	0.6708
		Hispanic	National	4	53	0.3104	0.3293
			State	-	100	0.6896	0.6708
		Other	National	4	162	0.2664	0.2375
			State	-	532	0.7336	0.7625
	Yes	Black	National	-	10	0.3713	0.3438
			State	-	23	0.6287	0.6562
		Hispanic	National	-	5	0.4250	0.3438
			State	-	6	0.5750	0.6562
		Other	National	-	38	0.3146	0.2554
			State	-	102	0.6854	0.7446
New York	No	Black	National	2	63	0.8139	0.7075
			State	4	17	0.1861	0.2925
		Hispanic	National	2	7	0.3582	0.7075
			State	4	11	0.6418	0.2925
		Other	National	2	153	0.3384	0.2994
			State	4	366	0.6616	0.7006
	Yes	Black	National	-	167	0.6774	0.5812
			State	-	108	0.3226	0.4188
		Hispanic	National	-	158	0.6395	0.5812
			State	-	79	0.3605	0.4188
		Other	National	-	136	0.3583	0.3766
			State	-	300	0.6417	0.6235
Ohio	No	Black	National	5	153	0.6183	0.4724
			State	6	119	0.3817	0.5277
		Hispanic	National	5	6	0.1810	0.4724
			State	6	24	0.8190	0.5277
		Other	National	5	308	0.3120	0.2583
			State	6	866	0.6880	0.7417
Pennsylvania	No	Black	National	4	24	0.4252	0.2555
			State	6	44	0.5748	0.7445
		Hispanic	National	4	1	0.0997	0.2555
			State	6	8	0.9003	0.7445
		Other	National	4	309	0.3526	0.3048
			State	6	704	0.6474	0.6952
	Yes	Black	National	-	51	0.5317	0.4881

Table 11-20. Comparison of compositing factors resulting from model to those used in the National Adult Literacy Survey — Continued

			State	-	61	0.4683	0.5119
		Hispanic	National	-	13	0.8522	0.4881
			State	-	2	0.1478	0.5119
		Other	National	-	75	0.3045	0.2693
			State	-	211	0.6955	0.7308
Texas	No	Black	National	4	77	0.5750	0.4069
			State	7	73	0.4250	0.5932
		Hispanic	National	4	204	0.4181	0.4069
			State	7	219	0.5819	0.5932
		Other	National	4	204	0.3450	0.3210
			State	7	477	0.6550	0.6790
	Yes	Black	National	-	123	0.6499	0.5185
			State	-	90	0.3501	0.4815
		Hispanic	National	-	87	0.5709	0.5185
			State	-	58	0.4291	0.4815
		Other	National	-	139	0.3507	0.3532
			State	-	317	0.6493	0.6468
Washington	No	Black	National	1	7	0.2327	0.1578
			State	5	31	0.7673	0.8422
		Hispanic	National	1	6	0.1368	0.1578
			State	5	33	0.8632	0.8422
		Other	National	1	99	0.1076	0.0821
			State	5	1055	0.8924	0.9179

demographic groups, uses other sample sizes for the state supplements, or includes elements other than those used in the 1992 designs, the compositing factors will have to be modified based on features of the new sample design. Equations (22) and (23) can be used to derive new factors. The population variances shown in earlier tables can be used to estimate the parameters in equations (22) and (23). If the oversampling rates are different from those in the 1992 National Adult Literacy Survey, the compositing factors will have to take into account the effect of variable sampling rates on the variance.

Hispanic respondents in the national sample. The oversampling introduced variability in the weights and increased the design effect for cases coming from the national sample. To best reflect this design feature, there should be separate compositing factors for minority and nonminority groups as in the first cycle of the National Adult Literacy Survey (refer to section 11.2.4 for details).

Chapter 12

CONSTRUCT VALIDITY OF THE ADULT LITERACY SCALES

Donald Rock and Kentaro Yamamoto, Educational Testing Service

The literacy scales were designed to measure adult literacy in three areas: prose literacy, document literacy, and quantitative literacy. The exercises in all three scales emphasize real-world problems or activities, although the prose scale includes some academic tasks. While all three scales are designed to measure an individual's literacy (and thus there is considerable overlap) each scale is designed to allow adult readers to demonstrate levels of literacy beyond those they might demonstrate on more focused academic reading exercises. The exercises that make up these three scales are designed to allow readers to take advantage of the unique, real-world experiences they bring to the testing situation. The prose scale comes closest to presenting passage materials in the form one would expect in the more traditional academic tests of reading performance. Nevertheless, the prose scale relies more heavily on expository rather than narrative materials—the kind of prose adults more typically read.

The development of the three literacy scales assumed that the level of literacy skill demonstrated in any assessment situation depends to some extent on the reader's knowledge of the passage content or familiarity with the mode of presentation of material (Kirsch and Mosenthal, 1990). For example, an auto mechanic who is familiar with technical manuals may show greater literacy skills on the document tasks than on the prose tasks. In short, performance on a reading test depends not only on comprehension of words and sentence structure but also on the content knowledge and experiences the reader brings to the testing situation. The adulf literacy scales were constructed to permit subpopulations with different backgrounds to take advantage of the specific skills that are unique to them and have these differences reflected in their profiles on the scales.

An important component of construct validity is the demonstration of discriminant validity. That is, scales that purport to measure different constructs should demonstrate significantly less than perfect inter-construct correlations (correlations between observed measures corrected for attenuation due to measurement errors). When one is attempting to disentangle measures that are expected by design to share considerable common variance, the problem becomes more complex. In the case at hand, the three literacy scales that were designed to measure prose, document, and quantitative literacy share many attributes. Still, the developers (Kirsch and Mosenthal, 1990) of the adult literacy scales presented both theoretical and pedagogical reasons for maintaining the three scales separately.

12.1 DATA AND MODELS

The purpose of this chapter is to estimate and test the discriminant validity of the three literacy scales from the perspective of correlation or covariance. Several models for the decomposition of the correlations among prose, document, and quantitative literacy items are proposed and compared. Two separate groups of data sets were used in the analysis described here.

12.1.1 Models for Four Restricted Adult Literacy Samples

The first group of data sets was extracted from the National Adult Literacy Survey. The total survey sample included about twenty-six thousand adults age 16 and older who were given one of 26 test booklets. Thirteen blocks of items were systematically spiraled within the 26 booklets, yielding a balanced incomplete block design. Each booklet contained three blocks, with each block containing items representing all three literary scales. Each test block contained from 12 to 15 items from each of the three scales—prose, document, and quantitative.

For this analysis, four booklets that contained no overlapping blocks were selected. The selection of booklets with no overlapping blocks allowed for a design that yielded independent estimates of the discriminant validity of the three scales across both people and items. The sample sizes for each of the four booklets selected were 831, 804, 784, and 804. These samples will be referred to as the restricted adult literacy samples.

The prose, document, and quantitative IRT scales assume an underlying unidimensional latent trait. Task responses to odd and even items within each selected booklet were summed into small testlets for the prose, document, and quantitative literacy scales. The scales were not necessarily designed to be parallel across test booklets, and some of the item parcels used to identify the first order factors had as few as six items (hence the term "testlet"). Scores were identified as the number of correct answers for each of the six testlets.

This chapter will supplement the classical factorial discriminant validity approach (Cronbach, 1984; Thorndike, 1982) by using structural equation models (Joreskog and Sorbom, 1979) to estimate a variance component decomposition of the three literacy scales as well as to model the relationship between selected demographic group membership variables and performance on the first order factors, while controlling for the general literacy factor. For one to argue that the first order factors have reliable unique variance and useful (valid) variance independent of general factor, one must show that the first order factors have some non-zero reliable unique variance, and that groups have different relationships with those factors.

Several structural equation models were used to explain variance in these testlets, which could take on different values in different samples. The simplest model is a first-order factor-analytic one:

$$y = \Lambda_y \eta + \varepsilon \tag{1}$$

where

y = a column vector consisting of six literacy testlets;

Λ_y = a matrix of factor loading of the testlets on the latent trait(s);

η = the first-order factor(s) or latent traits; and

ε = a column vector of error variances.

Depending on the number of first-order factors η, this model can represent either a single, general literacy factor η_L, or three separate literacy scales η_p, η_d, η_q. When three first-order factors are postulated, their intercorrelations are represented as the matrix Φ. If the three literacy scales do discriminate different latent traits, the values of these correlations should not be too close to unity.

It is also possible to decompose the first-order factors further, so that the first-order factors become functions of second-order factor(s) and exogenous variables.

$$\eta = B \xi + \Gamma X + \zeta \tag{2}$$

where

B = a matrix of factor loading of first-order latent traits on second-order factors;

ξ = the second-order factor; and

Γ = a matrix of regression coefficients of first-order latent traits on exogenous demographic and behavioral variables;

X = a vector of exogenous demographic and behavioral variables;

ζ = the unique, unexplained part of the first-order latent traits.

By substitution, the scores on the testlets can be represented in a single equation as

$$y = \Lambda_y (B \xi + \Gamma X + \zeta) + \varepsilon \tag{3}$$

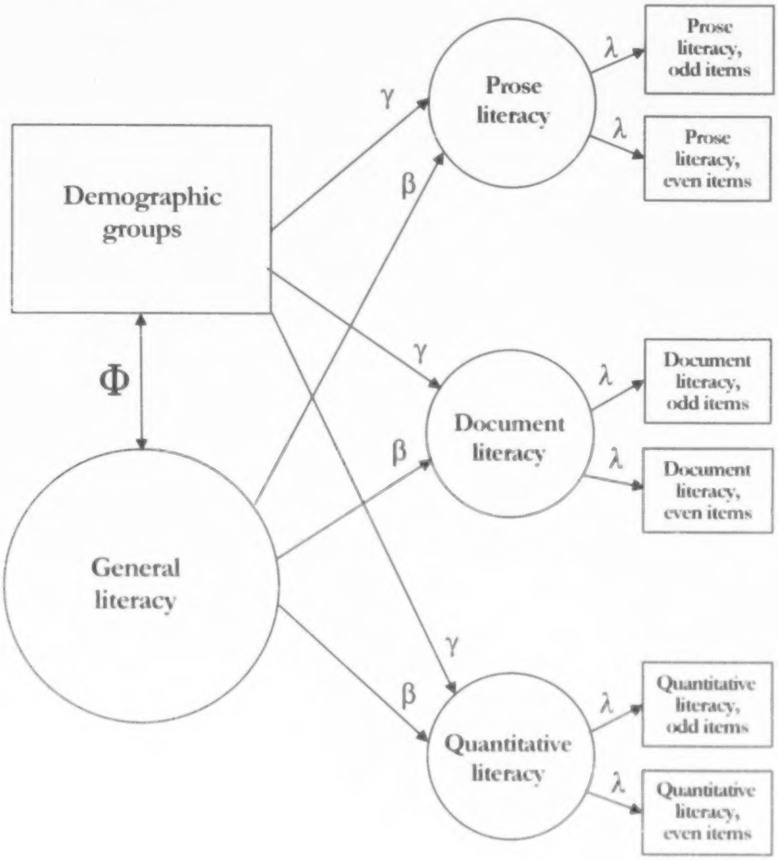
Discriminant validity is defined here as: (1) The variance decomposition of first order factors η_p, η_d, η_q into components related to the higher order factors ξ (such as general literacy), variance related to unique first-order factors ζ (the specific literacy scales), and error variance ε, and (2) group profile differences ΓX on first order factors η controlling for the general factor ξ_L. The term *higher order factor* refers to the fact that the three first order factors may share considerable variance, and the common skill ξ_L that underlies the three scales η_p, η_d, η_q is a general literacy skill. It is referred to as a higher-order factor because it is at a second level of summarization. One can view the first-order factors as summarizing the relationships among the observed scores, and the second-order factor as summarizing the relationships among the first order factors.

The reasons propounded by the scale developers (Kirsch and Mosenthal, 1990) for maintaining the three adult literacy scales as separate would imply that a single-factor model with a general literacy trait is not appropriate. The analysis of the restricted adult literacy samples investigates whether the three-factor model that differentiates prose, document, and quantitative literacy is consistent with the data, and if so, whether maintaining the three factors has any practical consequences. Joreskog and Sorbom (1993) discuss three types of situations with respect to fitting and testing models: 1) a strictly confirmatory situation, 2) alternative or competing models, and 3) a tentative initial model that is expected to be modified by the data, i.e., the final model is data driven. The present analysis is viewed as coming closest to the first type, a strictly confirmatory situation, where the three-factor model will be confirmed.

Figure 12-1 depicts the hypothesized relationships between the three first order factors, defined by the prose, document, and quantitative constructs, and a single higher order factor, called "general literacy." The identification of group differences on first order factors controlling for the general factor addresses the question about level differences on specific scales after partialling out the general literacy factor.

285

Figure 12-1. Model for discriminant validity based on subgroup differences on first-order factors


Demographic groups
General literacy
Φ
Prose literacy
Document literacy
Quantitative literacy
γ
β
λ
Prose literacy, odd items
Prose literacy, even items
Document literacy, odd items
Document literacy, even items
Quantitative literacy, odd items
Quantitative literacy, even items

12.1.2 Models for Seven GED/Adult Literacy Samples

The second group of samples was drawn from a larger sample of 1,577 General Educational Development (GED) test takers who took both the GED test battery and the adult literacy assessment as part of a larger study. The goals of this study were to examine the English-language literacy skills of a national sample of GED test-takers, including passers and non-passers, and to explore what the GED test and the National Adult Literacy Survey assessment measured in common and what was separately measured by each. Part of this project involved developing a prediction model relating scores on the GED scales to proficiencies on the adult literacy scales. Additional details on the study's design and findings were reported in Baldwin, Kirsch, Rock, and Yamamoto (1995).

While all individuals in this study took identical GED test booklets, seven assessment booklets from the National Adult Literacy Survey were spiraled through the sample. For analysis purposes this spiraled design was treated as seven independent sample replications. These seven replications will be referred to as the GED/adult literacy samples.

In a manner similar to the restricted adult literacy sample analysis, the GED scales, like the adult literacy scales, were split into "parcels" of items based on odd-even splits. The responses to odd and even items within the GED test were summed into testlets for the writing, social studies, science, literature, and mathematics constructs. Scores were identified as the number of correct answers for each of the ten testlets.

The GED/adult literacy samples permit partitioning the reliable variance of the adult literacy scales into two components—an academic, general (or "G") component and a scale-specific component. Like the models based on the four adult literacy samples, one can use the model to estimate group profiles on the first order factors independent of the general literacy factor. This type of group-related hierarchical model was originally developed by Muthen (1989) and subsequently applied to National Assessment of Educational Progress data (Muthen, 1994).

The use of the GED/adult literacy samples permitted investigation of related discriminant validity issues, using a model similar to that used for the restricted adult literacy samples analysis (except the exogenous variables were not used):

$$y = \Lambda_y \left(B \xi + \zeta \right) + \varepsilon \tag{4}$$

where

y = a column vector consisting of six literacy and ten GED testlets;

Λ_y = a matrix of factor loading of the testlets on the latent traits, either specific or general;

B = a matrix of factor loading of first-order latent traits (three literacy and five GED traits) on second-order factors;

ξ = the second-order factors (either one general factor for both literacy and GED scales, or two general factors, one for the GED and one for the literacy scales);

ζ = the unique, unexplained part of the first-order latent traits; and

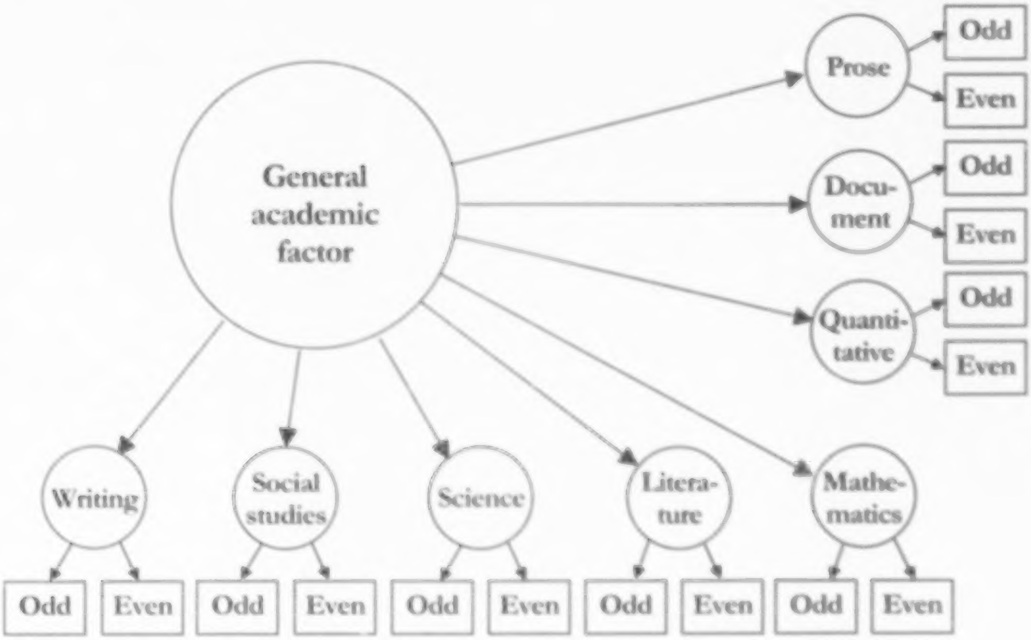
ε = a column vector of error variances.

As before, each member of a pair of testlets was hypothesized to mark a separate first order factor which in turn loads on one (or two) second order general factor(s).

In order to estimate the loadings of the adult literacy scales on an academic "G" factor, markers for such a "G" had to be measured along with the adult literacy scales. Since the GED/adult literacy samples included both the GED scales and the adult literacy scales, it was possible to estimate reliable unique variance for each of the ten scales that is uncorrelated with a single academic "G" factor that is hypothesized in the first model to underlie both the GED scales and the adult literacy scales.

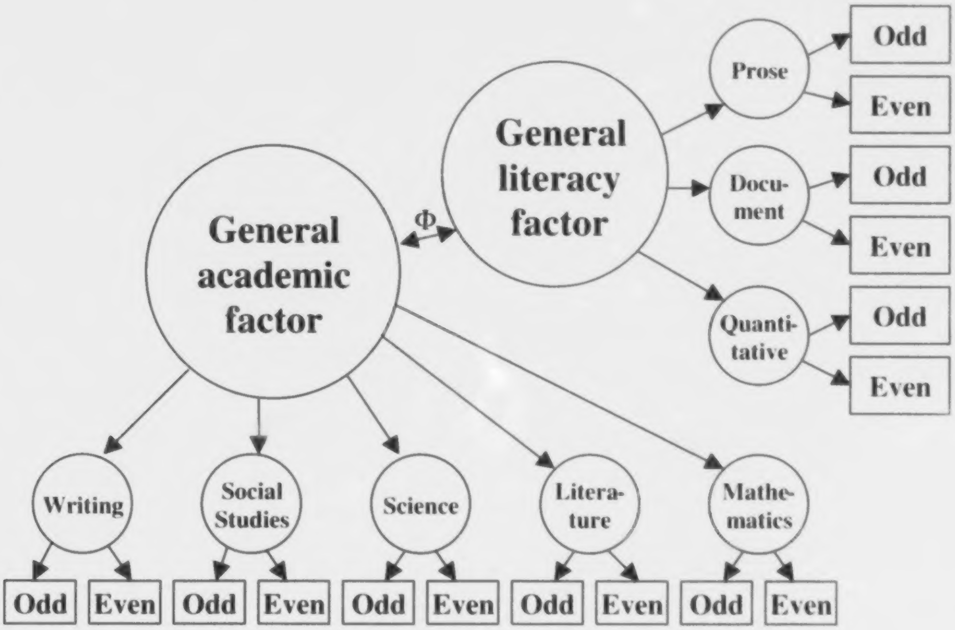
Figure 12-2 depicts the hypothesized relationships between the three first-order literacy factors, defined by the prose, document, and quantitative odd-even testlets; the five first-order GED factors, defined by the writing, social studies, science, literature, and mathematics odd-even testlets; and a single second-order factor, labeled a "general academic factor."

Figure 12-2. First hierarchical model based on seven GED/adult literacy samples


General academic factor
Prose
Odd
Even
Docu-ment
Odd
Even
Quanti-tative
Odd
Even
Writing
Social studies
Science
Litera-ture
Mathe-matics
Odd
Even
Odd
Even
Odd
Even
Odd
Even
Odd
Even

Under the second model, shown in Figure 12-3, while the literacy skills assessed in the three literacy scales have a shared general academic variance component, they also have significant non-zero components that are unique to each scale and, most importantly, are independent of the shared general academic variance component that underlies the GED scales. Figure 12-3 depicts the hypothesized relationships between the three first-order literacy factors and a second-order factor, called "general literacy factor" as well as between the first-order factors defined by the writing, literature, social studies, science, and mathematics GED scales and a single second-order factor, called "general academic factor." Since this model estimates separate second-order constructs underlying both the GED scales and the adult literacy scales, it was possible to estimate the unattenuated correlation between the adult literacy construct and the academic "G" construct as measured by the five GED achievement areas (writing, social studies, science, literature, and mathematics). The correlation between the two general second-order factors is labeled Φ.

Figure 12-3. Second hierarchical model based on seven GED/adult literacy sample


General literacy factor
General academic factor
Φ
Prose
Docu-ment
Quanti-tative
Odd
Even
Odd
Even
Odd
Even
Writing
Social Studies
Science
Litera-ture
Mathe-matics
Odd
Even
Odd
Even
Odd
Even
Odd
Even
Odd
Even

12.2 METHOD

The estimation of the model parameters was done using LISREL8 (Joreskog and Sorbom, 1992). A major advantage of the LISREL estimation of the hierarchical model is that it allows simultaneous estimation of both general and specific abilities defined by the eight GED and literacy scales. For the three literacy scales, the structural equation approach using second-order factors can partition the general and specific literacy abilities into orthogonal components and permits an assessment of the relative importance of the general literacy component and the three unique components of the scales. LISREL applies weighted least squares estimation procedures to the four matrices of polyserial correlations that were estimated separately in each of the four subpopulations, defined by which test booklet was taken. Polyserial correlations were used because the testlets had a restricted range of six or eight point scales, and the demographic variables ranged from dichotomies to four- or five-point scales.

Five variants of the model were developed for the four restricted adult literacy samples:

- The first variant was a very parsimonious, single-factor model for literacy, in which the factor loadings were constrained to be equal in each of the four samples.

- The second variant hypothesized three literacy factors (prose, document, and quantitative literacy), again with the same pattern of loadings in each sample.
- The third variant also hypothesized three literacy factors with the same pattern of loadings in each of the four samples, but this variant also constrained the intercorrelation among the factors to be equal in each of the four samples. This variant provides the best estimates of the factor intercorrelations.
- The fourth variant introduced a second-order factor underlying the three literacy scales and constrained the relationship between the general factor and the three first order factors to be the same across all four samples.
- The fifth variant also hypothesized three unique literacy factors and an underlying, second-order general literacy factor, but this variant also constrained all parameters except the parcel uniquenesses (errors) to be identical in all four samples.

The fifth variant was the most constrained model variant, with only the errors of measurement varying across populations. This solution provided a single summary variance component decomposition, which can be thought of as an average variance decomposition across the four samples.

Except for the last model variant, no constraints were put on the relationships between the observed indicators and their first order factors across samples, because the literacy tasks grouped together within booklets were not designed to be parallel across test booklets. Because the four subpopulations were random samples from the same total population, it is reasonable to estimate and test the assumption that the relationships among the latent variates are the same across the four random samples.

For all the final models that incorporate demographic or process variables, the relationship between the external variables (demographics) and the first order factors and the inter-relationships among the factors were assumed to be invariant across the four random samples. If there is no strong evidence to the contrary, then a single best estimate across all four replications of the relationship between the latent variates and the group dummy variables can be estimated.

Four variants of the model were developed for the seven GED/adult literacy samples.

- The first variant was again a parsimonious single-factor model for general academic ability underlying both the six literacy testlets and the ten GED testlets, in which the factor loadings were constrained to be equal in each of the seven samples.
- The second variant hypothesized three literacy factors (prose, document, and quantitative literacy), and five GED factors (writing, social studies, science, literature, and mathematics) underlying the odd-even pairs of testlets, again with the same pattern of loadings in each sample.
- The third variant introduced two second-order factors (academic "G" and general literacy) underlying both the five GED first-order factors and the three literacy first-order factors (Figure 12-3) and constrained the relationship between the general factors and the eight first-order factors to be the same across all seven samples.
- The fourth variant also hypothesized eight unique literacy factors, but constrained the model to a single underlying, second-order general factor (Figure 12-2), with all parameters except the parcel uniquenesses (errors) to be identical in all seven samples.

This solution represented the four variant component decomposition, or an average variance decomposition across the seven samples.

12.3 RESULTS AND DISCUSSION

12.3.1 Restricted Adult Literacy Samples

Table 12-1 presents various goodness of fit indices for five variant structural equation models across the four independent adult literacy samples of people and tasks. The non-normed fit index or Tucker-Lewis (1973) goodness of fit index indicates perfect fit on all models except the single factor model. Generally, non-normed fit indices of .95 or greater are considered acceptable fits.

The single-factor model variant does not fit the data well for virtually all the goodness of fit criteria shown in Table 12-1. This parsimonious single factor model has been included in the table as a kind of baseline model. However, it will be shown below that accepting such a parsimonious solution could result in masking group differences related to small but potentially identifiable specific components of variance associated with the three literacy scales.

The second and third model variants hypothesize three factors with the same pattern of loadings in each sample. While there is some additional lack of fit when going from the second to the third model (which adds the assumption that the intercorrelations among the factors are equal in the four samples), the root mean square error of approximation is considerably less than .05, which Browne and Cudeck (1993) consider a good fit. All other indices also suggest that the third model variant is consistent with the data.

Table 12-2 presents the factor intercorrelations and their standard errors (in parentheses) based on the third model variant. These inter-factor correlations are corrected for attenuation due to measurement errors. While the intercorrelations are relatively high, they are all significantly less than unity, given their standard errors.

Table 12-1. Goodness of fit indices for alternative adult literacy models with constraints across four restricted adult literacy samples

Model	Chi-square divided by degrees of freedom	Root mean square residual [a]	Non-normed fit index[b]	Root mean square error of approximation [c]
1. One first-order factor η_L, with same pattern of loadings $\Lambda_y^{(1)} = \Lambda_y^{(2)} = \Lambda_y^{(3)}$	10.28	0.032	.98	.054
2. Three first-order factors, η_p, η_d, η_q, with same pattern of loadings $\Lambda_y^{(1)} = \Lambda_y^{(2)} = \Lambda_y^{(3)}$	.97	0.006	1.00	.000
3. Three first-order factors η, same pattern and same correlations among factors $\Lambda_y^{(1)} = \Lambda_y^{(2)} = \Lambda_y^{(3)}$ $\Phi^{(1)} = \Phi^{(2)} = \Phi^{(3)}$	3.00	0.019	1.00	.025
4. Three first-order factors η, one second-order factor ξ, with $\beta^{(1)} = \beta^{(2)} = \beta^{(3)}$	1.68	0.015	1.00	.015
5. Three first-order factors η, one second-order factor ξ, with $\Lambda_y^{(1)} = \Lambda_y^{(2)} = \Lambda_y^{(3)}$ $\beta^{(1)} = \beta^{(2)} = \beta^{(3)}$	5.6	0.044	.99	.038

[a] The root mean square residual can be loosely interpreted as the average partial correlations among the observed variables after the model is fitted.

[b] The Tucker-Lewis measure of goodness of fit of the model to the data. A measure of the "reliability" of the model that is essentially independent of sample size.

[c] The root mean square error of approximation is a measure of the discrepancy between the sample and population estimates adjusted for the degrees of freedom.

Table 12-2. Factor intercorrelations (and standard errors) based on third model

	Prose literacy		Document literacy		Quantitative literacy	
Prose literacy	-	-				
Document literacy	0.91	(0.00)	-	-		
Quantitative literacy	0.87	(0.00)	0.88	(0.01)	-	-

The fourth and fifth model variants estimate the hierarchical solution and constrain the relationship between the general factor and the three first order factors to be the same across all four samples. The fifth model differs in that it constrains all but the errors of measurement to be the same across populations. Table 12-3 shows the average estimates of the variance components based on this solution. The entries in the table can be interpreted as percentages. That is, the entries give the percentage of each variable's variance that is attributable to a general literacy factor, a scale-specific factor, and error. For example, the variance of the odd items of the document scale is decomposed into 70 percent attributable to general literacy, 5 percent to scale unique variance, and 25 percent to error.

Table 12-3. Components of variance decomposition of the three literacy scales based on fifth model

Test parcels	General literacy variance (the common part, or $\hat{\Lambda}_y \hat{\Gamma} \hat{\Phi} \hat{\Gamma}' \hat{\Lambda}_y'$)	Scale-specific literacy variance (the specific part, or $\hat{\Lambda}_y \hat{\Psi} \hat{\Lambda}_y'$)	Error variance
Prose literacy, odd items	.623	.078	.299
Prose literacy, even items	.654	.081	.265
Document literacy, odd items	.698	.053	.249
Document literacy, even items	.730	.005	.215
Quantitative literacy, odd items	.641	.132	.227
Quantitative literacy, even items	.671	.138	.191

The maximum likelihood estimates of the reliability of the three scales based on the hierarchical solution (model four) are .80, .85, and .86, for the prose, document, and quantitative scales respectively. These estimates do not assume that the odd and even splits are necessarily parallel within a population. Inspection of Table 12-3 indicates that the quantitative scale has almost twice as much specific variance as either the prose or document scale. The magnitudes of the specific components suggest that the size of the scale-specific components is sufficient that they may be useful in identifying group differences independent of the general common literacy factor.

The possibility that the specific variances associated with each scale are sufficiently unique to yield different profiles of subgroup means is examined through regressions of first order factors on group indicators, while partialling out the general factor. Table 12-4 presents the regressions of the first order factors on the group dummy variables controlling for the general literacy factor (the second order factor).

Table 12-4. Regression coefficients (and standard errors) of first-order factors—the three literacy scales—on selected demographics and literate activities, holding constant the general literacy second-order factor

Independent variables	Prose factor	Document factor	Quantitative factor	Correlations between Demographics and General Factor
Demographics				
Male	-0.03 (.02)	-0.02 (.02)	0.04*(.02)	.02
Under 55 years old	0.01 (.02)	-0.04* (.03)	0.04*(.02)	-.08
White vs. other	-0.01 (.02)	-0.01 (.02)	0.02 (.02)	.46
Working full-time	0.03 (.02)	-0.04*(.02)	0.03 (.02)	.35
College graduate	0.04*(.02)	0.07*(.02)	0.02 (.02)	.53
Literate activities				
Reading newspapers	0.03 (.02)	-0.01 (.02)	0.00 (.02)	.19

Note: The standard errors of the regression estimates are shown in parenthesis. An asterisk next to a regression coefficient indicates that the estimate was significantly different from zero.

In the case of gender, males do better than females on the quantitative scale when controlling for general literacy, while there is no difference between the sexes on the prose or document scale. This finding is, of course, consistent with other studies of gender differences on quantitative material, which do not typically control for general literacy ability.

Other demographic groupings that show differential performance on the first order factors while controlling for general literacy are:

- *Age*: Young individuals tend to do worse on the document literacy and better on the quantitative scale than older adults. It may be that the older population is more familiar with some of the tasks in the document scale. The document scale probably has the most real-world orientation of the three scales and is the least sensitive to formal learning.
- *Race/Ethnicity*: There were no differences between whites and all other racial/ethnic groups combined on the three literacy scales after controlling for the general literacy factor (see Table 12-4). Sample size limitations prevented treating racial and ethnic minorities separately.
- *Employment*: Individuals who were working full time had significantly lower scores on the Document scale then those were either unemployed or underemployed when controlling for general literacy.
- *Education*: The more highly educated individuals tended to do better on both the prose literacy scale and worse on the document literacy scale when the general literacy factor is controlled. This finding is consistent with the observation that the tasks in the prose scale are more similar to those given in academic situations while those in the document scale have a more real-world orientation.

None of the relationships between newspaper reading activities and the first order factors were significant.

12.3.2 GED/Adult Literacy Samples

Table 12-5 presents goodness of fit statistics for four variant structural equation models estimated on the seven GED/adult literacy samples of people and tasks. The non-normed fit index or Tucker-Lewis (1973) goodness of fit index indicates excellent fit on all models except the single factor model, which is primarily a baseline model that could mask group differences related to specific components of variance associated with the three literacy scales.

The second model variant hypothesizes eight factors with the same pattern of loadings in each sample. The goodness of fit statistics for this model provide the estimates of the factor intercorrelations presented in Table 12-6. These inter-factor correlations are corrected for attenuation due to measurement errors. While the intercorrelations are relatively high, they are all significantly less than unity, given their standard errors. The factor intercorrelations among the adult literacy scales are significantly lower in the GED/adult literacy samples than those found in the restricted adult literacy samples. The lower intercorrelations reflect the more homogenous population found among the GED test takers. It should be noted that among the adult literacy scales the document scale has consistently lower correlations with all the GED scales.

Table 12-5. Goodness of fit indices for alternative adult literacy models with constraints across seven GED/adult literacy samples

Model	Chi-square divided by degrees of freedom	Root mean square residual [a]	Non-normed fit index[b]	Root mean square error of approximation [c]
1. Single factor	0.16	0.090	.78	.060
2. Eight factor (five GED, three adult literacy) $\underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_7$	1.42	0.070	.98	.020
3. Hierarchical model, with two general factors (GED and adult literacy) and eight specific factors (scales), with $\underline{\gamma}_1 = \underline{\gamma}_2 = \ldots = \underline{\gamma}_7$ $\underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_7$	1.96	0.070	.96	.03
4. Hierarchical model, with one general factors (academic "G") and eight specific factors (scales) $\underline{\gamma}_1 = \underline{\gamma}_2 = \ldots = \underline{\gamma}_7$ $\underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_7$ $\underline{\Lambda}_{y1} = \underline{\Lambda}_{y2} = \ldots = \underline{\Lambda}_{y7}$	2.52	0.090	.94	.030

[a] The root mean square residual can be loosely interpreted as the average partial correlations among the observed variables after the model is fitted.

[b] The Tucker-Lewis measure of goodness of fit of the model to the data. A measure of the "reliability" of the model that is essentially independent of sample size.

[c] The root mean square error of approximation is a measure of the discrepancy between the sample and population estimates adjusted for the degrees of freedom.

The third model in Table 12-5 provides an estimate of the correlations between the two second-order factors. The first of two second-order factors is an estimate of an academic "G" factor based on the shared variance among the five GED scales, and the second is an estimate of the general literacy factor that underlies the correlations among prose, document, and quantitative literacy. This inter-construct correlation, labeled Φ in the structural equations modeling literature, was .78 with a standard error of .01. While the two constructs are highly correlated, they seem to be measuring somewhat different things.

Table 12-6. Factor intercorrelations (and standard errors) among the GED and literacy scales corrected for attenuation

	Writing	Soc Stud	Science	Lit	Math	Prose	Document
GED scales							
Writing	1.00 (-.--)						
Social Studies	0.77 (0.01)	1.00 (-.--)					
Science	0.77 (0.01)	0.92 (0.01)	1.00 (-.--)				
Literature	0.79 (0.01)	0.91 (0.01)	0.86 (0.01)	1.00 (-.--)			
Math	0.68 (0.02)	0.77 (0.01)	0.80 (0.01)	0.67 (0.02)	1.00 (-.--)		
Literacy scales							
Prose	0.62 (0.02)	0.70 (0.02)	0.67 (0.02)	0.71 (0.02)	0.60 (0.02)	1.00 (-.--)	
Document	0.56 (0.02)	0.56 (0.02)	0.56 (0.02)	0.57 (0.02)	0.56 (0.02)	0.81 (0.02)	1.00 (-.--)
Quantitative	0.59 (0.02)	0.69 (0.02)	0.70 (0.02)	0.61 (0.02)	0.81 (0.02)	0.82 (0.02)	0.77 (0.02)

The fourth model variant in Table 12-5 (corresponding to Figure 12-2) estimates the single-factor hierarchical solution while constraining the relationships between the general factor and the eight first-order factors, as well as all but the errors of measurement to be the same across all seven samples. This solution provided the components of the variance decomposition averaged over the seven GED/adult literacy samples shown in Table 12-7. The entries in this table give the percentage of each variable's variance that is attributed to the general factor, to the scale specific factor, and to error. For example, about 40 percent [26/(38+26)] of the prose literacy odd testlet's reliable variance is independent of the general academic "G" factor, while 56 percent [35/(35+27)] of the document literacy odd testlet's reliable variance is independent of academic "G." Among the GED scales only writing and mathematics had non-trivial reliable components of variance that were independent of the general factor. This suggests that the academic "G" is probably better described as an academic verbal "G." The fact that the document scale had a proportionately greater component of variance independent of the general factor than the other adult literacy scales is consistent with the observation that the prose and quantitative scales are more likely to have tasks similar to those found in their counterpart academic scales.

Table 12-7. Components of variance decomposition of GED and literacy scales into academic "G," specific, and error components (in percentages)

Scale Parcels (Testlets)	Reliable Variance: Common (Academic G)	Reliable Variance: Specific (Scales)	Error
GED scales			
Writing (odd)	46	23	31
Writing (even)	57	28	15
Social Std. (odd)	80	07	13
Social Std. (even)	76	07	17
Science (odd)	73	09	18
Science (even)	73	09	18
Literature (even)	69	13	18
Literature (odd)	67	13	20
Math (odd)	53	26	21
Math (even)	57	28	15
Adult literacy scales			
Prose (odd)	38	26	36
Prose (even)	34	24	42
Docum. (odd)	27	35	38
Docum. (even)	23	30	47
Quant. (odd)	35	24	41
Quant. (even)	37	25	38

12.4 CONCLUSIONS

The three literacy scales—prose, document, and quantitative literacy—are relatively highly related, as indicated by their factor intercorrelations. However, even with disattenuated correlations in the high eighties to the low nineties, there is still room for some differences in group profiles due to small, but reliable specific components of variance in the three literacy scales. While these differences are relatively small, their direction seems to be consistent with expectations about the different contents of the scales.

The correlation between the higher order factor underlying the GED scales and its counterpart underlying the literacy scales was .78, which is close to the same order as that commonly found between verbal and quantitative measures. This suggests that the common factor underlying adult literacy is related to academic "G" but is not the same thing. Among the adult literacy scales, the document scale shared the least variance with the academic "G" factor. About 40 percent of the reliable variance of the prose and quantitative scales was independent of the general academic factor, while about 56 percent of the document scale was independent of academic "G."

Chapter 13

INTERPRETING THE ADULT LITERACY SCALES AND LITERACY LEVELS

Irwin Kirsch and Ann Jungeblut, Education Testing Service;
Peter Mosenthal, Syracuse University

The major benefits resulting from statistically derived scales are the enhancement of the comparability of results across groups, age, and time, and the provision of a basis for relating background and attitude variables to performance (Messick, Beaton, and Lord, 1983). But however useful such statistically derived scales may be, a need remains to provide supplementary information aimed at guiding the interpretation of the scales. This chapter outlines chronologically the development of a theory of task characteristics that contribute to task difficulty on the three adult literacy scales. This chapter also explains the reasoning that provided the basis for developing five literacy levels on each scale and the process that led to general descriptions of the task characteristics associated with these literacy levels.

13.1 BACKGROUND

Historically, standardized objective tests have provided a means of comparing the performance of one individual to that of a relevant group or groups. A robust theory of mental tests and measurement, including statistical theory and procedures, was developed over the years to facilitate appropriate test interpretation and use. Such techniques have served well in both individual assessments and in large scale surveys. The shift of interest to criterion-referenced testing has yielded more information on task difficulty and the percentages of people able to perform certain tasks. What has been missing is a means of looking at the interaction between task characteristics and people's performance in order to enhance both the interpretability and utility of test results, as well as decisions and actions based on test scores.

The display of both people and tasks along a common scale invites the question of whether tasks receiving similar score values share certain characteristics. This display also raises the issue of to what extent these characteristics (as well as the response consistencies of individuals) differ systematically from one end of the scale to the other. Some important benefits that derive from the systematic exploration of these issues are that they:

- increase understanding of variables that contribute to task difficulty;
- enhance the ability to generate new tasks that more fully represent the domain(s) being assessed;
- establish a context in which one can define the domain boundaries, that is, enhance score meaning; and
- strengthen the links among testing, research, practice, and policy.

Collectively, these benefits contribute to an improved theoretical framework that systematically helps to account for consistency in task responses. Rather than treating the task responses as a conglomeration of specifics, these response consistencies are typically summarized in the form of scores or sub-scores. Although discrete behaviors and isolated observations may be of interest, in terms of measurement validity they are far less meaningful and dependable than response consistencies (Messick, 1989).

The purpose of this chapter is to trace the evolution of the theoretical framework used to construct, interpret, and report large-scale literacy survey data. Prior to the 1992 National Adult Literacy Survey, two other assessments were conducted that used similar methods—the 1985 young adult literacy assessment and the 1990 survey of the literacy of job-seekers served by the U.S. Department of Labor. Together, these three surveys have employed a common definition of literacy, thereby contributing to the evolution of a rich theoretical framework for literacy assessment. The 1992 survey included each of the 1985 literacy blocks in the 1992 cognitive instrument. This chapter describes each study in terms of its contributions to the expanding theoretical understanding of literacy, along with its practical application to literacy measurement, including the development of five proficiency levels used to interpret and report the 1990 and 1992 survey results. The issues and empirical evidence presented address various aspects of validity.

13.2 THE 1985 YOUNG ADULT LITERACY ASSESSMENT

The 1985 young adult literacy assessment was funded with a Federal grant under the National Assessment of Educational Progress (NAEP) program. This survey was designed to assess the literacy skills of young adults 21 to 25 years old. The deliberations of the expert panels that oversaw the development of the young adult literacy assessment led to the adoption of the following definition of literacy: *Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential* (Kirsch and Jungeblut, 1986a). Reflecting this definition, the organizing theoretical framework that evolved for task development in this study was a multidimensional approach to literacy—that is, tasks were developed to cover the three distinct areas of prose, document, and quantitative literacy.

Literacy tasks for the young adult assessment were based on the printed *material* to be read and the *purpose* which the reader brought to the material. "Material" refers to the linguistic form in which the information is displayed. Twelve categories of material were identified: sign/label, directions, memo/letter, form, table, graph, prose, index/reference, notice, schematic or diagram, advertisement, and bill/invoice. "Purpose" refers to why the reader engages in the task, or what information the reader is seeking. The reader's purpose influences both the strategies and cognitive operations in which the reader engages while completing the task. Five categories of purpose, reflecting various levels of processing were identified:

knowledge, evaluation, specific information, social interaction, and application. The five categories were defined as follows:

- Knowledge—reading to integrate information, to remember sets of facts for later use, or to go beyond information given;
- Evaluation—comparing and contrasting points of view or using printed information to make a reasoned judgment;
- Specific information—locating a specific fact to satisfy a particular need, such as looking up a fact in a reference book;
- Social interaction—organizing and sequencing information to communicate to another person or group, such as preparing a memo, or writing a letter; and
- Application—following instructions to construct, make, or repair something, doing simple calculations, or providing simple facts to complete forms.

Crossing the twelve identified materials with the five identified uses resulted in the matrix shown below in Table 13-1, taken from Kirsch and Jungeblut (1986b). The dots in the table indicate cells for which literacy tasks were developed and organized into blocks for administration in the 1985 assessment. In developing tasks, primary emphasis was placed on representing the broad range of literacy behaviors people frequently encounter in occupational, social, and educational settings (Guthrie, Siefert, and Kirsch, 1986). It was felt that simulations of the skills in context rather than traditional multiple-choice questions would provide a more ecologically valid and useful assessment of literacy competencies. Efforts were undertaken to create assessment materials that would address these concerns. For example, the 1985 assessment printed a 4-page newspaper containing a selection of articles that had appeared in national newspapers. Respondents were asked to summarize arguments from an editorial, to locate specific information in a news story, and to look up information in a TV listing or a classified page.

Information on the additional considerations underlying the development of the new literacy tasks for the 1992 survey can be found in Campbell, Kirsch, and Kolstad (1992) *Assessing Literacy: The Framework for the National Adult Literacy Survey* and in Chapter 4 of this report.

Table 13-1. Matrix of materials and uses for adult literacy tasks

Materials	Uses				
	Knowledge	Evaluation	Specific information	Social interaction	Application
Sign/Label			✓		✓
Directions					✓
Memo/Letter				✓	
Form		✓	✓		✓
Table		✓	✓	✓	✓
Graph		✓	✓		
Prose	✓	✓	✓	✓	
Index/Reference			✓		
Notice		✓	✓		✓
Schematic or Diagram			✓		✓
Advertisement			✓		✓
Bill/Invoice			✓		

The intersection of the linguistic form in which information is displayed (materials) and the type of information needed or sought (use) in this table not only provided the operational definition of a literacy task but also determined the information-processing demands required for successful performance. The approach to literacy task development guiding the study led to the anticipation of more than a single dimension of literacy. Although the number of cells was too small to allow finding a separate dimension for each of the filled cells in Table 13-1, the designers of the 1985 assessment explored a number of plausible alternative organizing structures on an *a priori* basis before the 1985 data were available for analysis. It is the interaction of the materials and purposes that was expected to define task difficulty and, thus, the placement of tasks on the literacy scales.

13.2.1 Dimensionality of Literacy Skills

Prior to the 1985 assessment, there had been a marked tendency to describe literacy in terms of the ability to perform successfully a series of concrete tasks, e.g., to complete an application for a driver's license, to comprehend the warning on a container of poison, and to interpret familiar street signs (Murphy, 1973). For the most part, success was summed across such diverse tasks and an arbitrary cutting point established (e.g., 75 percent correct), below which an individual is classified as "functionally illiterate" (NAEP, 1976). Such an approach, with its lack of an organizing principle and arbitrary cutpoint used, was dismissed for the 1985 assessment, since it would contribute nothing to understanding the process of literacy. Moreover, this approach was in direct conflict with the theoretical framework of the 1985 study (Kirsch and Jungeblut, with others, 1986, p. III-2).

Previous theoretical approaches to the study of literacy had used a variety of organizing principles. According to one traditional approach, literacy skills had been categorized into reading, writing, speaking, listening, and arithmetic or mathematics. Another approach was to organize disparate literacy tasks by the context in which they occur: home, school, or work. Yet another approach was to categorize literacy tasks in terms of the materials or formats in which they occur and to examine the associated types of purposes both within and across materials. As an instance reflecting a similar distinction, the 1972 NAEP reading assessment came to aggregate reading exercises in terms of "themes"-word meanings, visual aids, written directions, references materials, significant facts, main ideas, inferences, and critical reading (Kirsch and Jungeblut, with others, 1986, p. III-3).

The 1985 assessment designers reached a compromise among the various organizing concepts and hypothesized three scales: a prose literacy scale, a document literacy scale, and a quantitative literacy scale. In this way, they were able to acknowledge that the mental processes underlying proficiency with prose texts are probably qualitatively different from those underlying proficiency with documents and that both are distinct from proficiency with quantities that can be embedded in either prose texts or documents (Kirsch and Jungeblut, with others, 1986, p. III-4).

A factor analysis was performed on the 1985 data in order to explore dimensionality. This was done to find evidence in the empirical data to substantiate the three literacy scales. The product-moment correlation coefficients among the tasks, with squared multiple correlations inserted as communality estimates in the main diagonal, were factor analyzed by the method of principal axes. The mean squared multiple correlation was 0.92 (trace = 101.01). An examination of the latent roots revealed three sizable factors followed by several smaller factors (roots = 18.11, 2.89, 2.30, 2.00, 1.94, 1.87, 1.79, 1.68, 1.67, 1.58,...). Following the logic of Cattell's (1966) *scree* test, the breaks in the pattern of latent roots indicated at least three salient factors with the possibility of at least five additional factors. Analysis of parallel random data reinforced the judgment that a three-factor solution was appropriate. However, for exploratory purposes three separate analyses were conducted: one in which eight factors were retained and rotated for interpretation; another in which five factors were retained; and, a third in which, three factors were retained for rotation and interpretation (Kirsch and Jungeblut, with others, 1986, p. III-5).

In each instance, the factors were rotated to simple structure by the varimax procedure and to oblique simple structure by the DAPPER method (Tucker and Finkbeiner, 1981). The DAPPER method was selected specifically to allow the complex literacy tasks to load on more than one factor. Indeed, many of the literacy tasks did so (Kirsch and Jungeblut, with others, 1986, p. III-6). Tasks loading highest on the first and largest factor seemed to rely heavily on prose comprehension; tasks loading highest on the second factor seemed to reflect skill in using documents, while those tasks loading highest on the third factor required the application of arithmetic operations. The DAPPER method was selected specifically to allow the complex literacy tasks to load on more than one factor.

The intercorrelations among the literacy scales for the total group of 3474 respondents ranged from .49 to .56, thus revealing only a moderate level of association. Indeed, the intercorrelations provide further support for the notion that literacy skills can and should be separated along at least three distinct dimensions—prose, document, and quantitative skills. These important distinctions would be lost if the diverse tasks from the 1985 young adult literacy assessment had been aggregated and reported on a single scale (Kirsch and Jungeblut, with others, 1986, p. III-34).

13.2.2 Difficulty of Literacy Tasks

Since the 1985 young adult literacy assessment was funded under the NAEP program, it tended to adopt many of the survey design and statistical methods used in that program. Beginning with 1983-84 reading assessment of school children, NAEP chose to anchor items representing standard deviation units along the reading proficiency scale. The exemplar items selected discriminated between each pair of standard deviation units in the following way: The NAEP reading proficiency scale was designed to extend from 0 to 500 with a mean of 250 and a standard deviation of 50. Thus, the selected anchor points were 150, 200, 250, 300, and 350 (Beaton and Allen, 1992). The criteria for selecting exemplars at each anchor point were that 80 percent or more of the students at that point (e.g., 250) answered the item correctly, while less than 50 percent of the students at the next lower level (e.g., 200) answered the same item correctly (Kirsch and Jungeblut, with others, 1986, p. III-9). The exemplar items identified through this procedure represented advances in student reading proficiency from one anchor point to the next. A panel of content experts then examined the exemplar items near each anchor point and identified the reading knowledge, skills, and abilities demonstrated by students answering each item correctly. The panelists' descriptions were summarized to characterize performance at each anchor level (Phillips, et al., 1993).

In the 1985 assessment, the three literacy scales were designed to range from 0 to 500, with a mean of 250 and a standard deviation of 50 points.[1] Unlike the 1983-84 NAEP reading assessment, the 1985 young adult literacy assessment had relatively few literacy tasks on the prose and quantitative literacy scales, so it was not feasible to provide meaningful descriptions at identical numerical points (e.g., standard deviation units) on each of the three scales. In addition, one would not expect that on each of the scales, tasks exemplifying important shifts in their cognitive demands would fall at comparable points (Kirsch and Jungeblut, with others, 1986, p. III-9).

---

[1] The scales used in 1992 were linked to the 1985 scales, so the mean could not be fixed at 250. The IRT models that structure these scales are described in Chapter 9.

To help guide interpretation of the three literacy scales and to attempt to identify factors associated with task difficulty, an item mapping procedure reflecting response probabilities was employed. Following NAEP's anchoring system at the time, tasks were placed on the scale at the point at which a minimum of 80 percent of the young adults at a particular ability level could be expected to complete the tasks successfully (see Chapter 14 for additional information). An additional criterion for selection of exemplar tasks was that approximately 30 percentage points (in terms of IRT response probabilities) separated individuals scoring around the same scale value as the task placement from individuals scoring one standard deviation (i.e., 50 points) lower. In the context of the 1985 young adult literacy assessment, the particular exemplars not only met this anchoring criterion representing performance at various levels of difficulty but also were seen to reflect a combination of increasingly complex skills interpreted to be associated with successful performance as task difficulty increased.

Once tasks were located on the scales, it was possible to summarize survey results by presenting selected exemplar tasks placed around successive points on the three scales along with percentages of people scoring at or above those same points. It was anticipated that such a graphic presentation would help give meaning to the scale and, thereby, increase the interpretability of results for the total group, as well as for subgroups of particular interest.

The next three sections of this chapter describe the mapping of task difficulty, the proficiencies of young adults on the 1985 prose, document, and quantitative scales and the identification of characteristics that underlie task difficulty at various points on the three scales. Task characteristics were identified on the basis of the complexity of the information-processing demands required for successful performance, rather than by features of the text alone, such as vocabulary or sentence length.

13.2.3 Prose Comprehension Scale

Exhibit 13-1 presents information about task difficulty and population performance on the prose literacy scale based on the item map reported from the young adult assessment (Kirsch and Jungeblut, 1986a). It can be readily seen, for example, that 90 percent or more of young adults in America demonstrated proficiency on the tasks extending down from 225 on the prose literacy scale. The tasks associated with points below 225 in the column on the left range from skill in locating one feature of information in a sports article to writing about a job one would like. In the original report, the full figure (not reproduced here) showed results not only for the total population but also for various racial/ethnic groups and for different levels of educational attainment. While each subgroup differed in their mastery at various levels on the scale, the ordering of the task difficulties remained the same across all groups.

Exhibit 13-1. Percentages of adults and selected tasks at or above successive points on the prose literacy scale: Adults 21 to 25 years old, 1985

	Selected tasks at decreasing levels of difficulty*	Selected points on the scale	Percent of total
		500	
397	Identify appropriate information in lengthy newspaper column		
387	Generate unfamiliar theme from short poem		
		375	8.8 (0.7)
371	Orally interpret distinctions between two types of employee benefits		
361	Select inappropriate title based on interpretation of news article		
		350	21.1 (1.1)
340	State in writing argument made in lengthy newspaper column		
339	Orally interpret a lengthy feature story in newspaper		
		325	37.1 (1.6)
313	Locate information in a news article		
		300	56.4 (1.5)
281	Locate information on a page of text in an almanac (3-feature)		
279	Interpret instructions from an appliance warranty		
278	Generate familiar theme of poem		
277	Write letter to state that an error has been made in billing		
		275	71.5 (1.4)
262	Locate information in sports article (2-feature)		
		250	82.7 (1.2)
		225	90.8 (0.7)
210	Locate information in sports article (1-feature)		
199	Write about a job one would like	200	96.1 (0.5)
		0	

*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

The 15 tasks comprising the prose comprehension scale appeared to reflect three qualitatively different aspects of reading comprehension: 1) matching of literal and corresponding information; 2) producing and interpreting text information; and 3) generating a theme or organizing principle from text information. Each of these three aspects contributed to a broad range of difficulty, with significant overlap among the three. Exhibit 13-2 presents information about task difficulty for the three aspects of the prose comprehension scale, again based on an item map reported in the young adult assessment (Kirsch and Jungeblut, 1986a).

Proficiency in Matching Literal Corresponding Information represents a continuum defined by the number of features that readers must identify to match information asked for in a question or directive with explicit or corresponding information in the text. At the simplest end of the continuum, readers match requested information with information in the text on the basis of a single, commonly shared feature. At the middle and upper ends, the match involves several features or several categories of information.

Proficiency in Producing and Interpreting Text requires readers to use background knowledge or textual information. Response at the simplest level involves producing personal background information. At more difficult levels, readers may have to interpret the directive or compare and contrast information.

Proficiency in Generating a Theme or Organizing Principle from Text Information requires readers to synthesize information consistent with arguments in the text. At the simplest level, the reader's task is to generate a theme from relatively short text. Generating the theme becomes more difficult as the concept becomes less familiar, or the arguments are less repetitive or more widely separated in lengthy text.

13.2.4 Document Literacy Scale

Exhibit 13-3 reproduces information about task difficulty and population performance on the document literacy scale based on the original item maps from the young adult literacy assessment (Kirsch and Jungeblut, 1986a). Again, it can be seen that 90 percent or more of the total group of young adults surveyed demonstrated proficiency on the tasks extending down from 225 on the document literacy scale. The tasks associated with points below 225 in the column on the left include entering the date on a bank deposit slip, identifying the cost of a particular theater trip from among those given in a notice, entering personal information on a job application form, locating the expiration date on a driver's license, and signing one's name on an image of a Social Security card.

The 43 tasks comprising the 1985 document literacy scale begin with a question or directive. The reader must first identify the important information in the question to be matched to information in a document. Among the most important characteristics associated with task difficulty are: 1) the number of features that readers must identify in a question or directive and match with features of information in a

Exhibit 13-2. Selected tasks and corresponding levels of difficulty* defining the three aspects of the prose comprehension scale: Adults 21 to 25 years old, 1985

	Matching Literal and Corresponding Information	Producing and Interpreting Text	Generating a Theme	Selected points on the scale
				500
397	3-feature match from newspaper article (corresponding)			
387			Generate theme from single unfamiliar metaphor	
				375
371		Interpret job-related benefit classification		
				350
340			Generate theme from repetitive argument widely dispersed	
				325
				300
281	3-feature match from a page of text in an almanac (literal)			
279		Interpret appliance warranty		
278			Generate familiar theme from argument	
				275
				250
				225
210	1-feature match from newspaper article (corresponding)			
199		Produce text using personal background information		200
				0

*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

Exhibit 13-3. Percentages of adults and selected tasks at or above successive points on the document literacy scale: Adults 21 to 25 years old, 1985

	Selected tasks at decreasing levels of difficulty*	Selected points on the scale	Percent of total
		500	
[illegible]	Use bus schedule to select appropriate bus for given departure		
320	Use sandpaper chart to locate appropriate grade given specifications	325	37.6 (1.6)
300	Follow directions to travel from one location to another using a map	300	57.2 (1.7)
294	Identify information from graph depicting source of energy and year		
278	Use index from an almanac	275	73.1 (1.2)
262	Locate eligibility from table of employee benefits		
257	Locate gross pay-to-date on pay stub		
255	Complete a check given information on a bill		
249	Locate intersection on street map	250	83.8 (1.0)
221	Enter date on a deposit slip	225	91.0 (0.8)
219	Identify cost of theatre trip from notice		
211	Match items on shopping list to coupons		
196	Enter personal information on job application	200	95.5 (0.5)
192	Locate movie in TV listing in newspaper		
181	Enter caller's number on phone message form		
169	Locate time of meeting on a form	175	98.4 (0.3)
160	Locate expiration date on driver's license		
		150	99.7 (0.1)
110	Sign your name	0	

*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

document; 2) the degree to which feature information given in the question or directive corresponds to, or is closely identified with, the requested information in the document; and 3) the number of exemplars or representations in the document that have at least one feature in common with those in the question, thereby serving as distractors or plausible correct answers for the reader.

Once a match between a question (or directive) and document information is made, the reader must determine whether the information matched is sufficient. If it is insufficient, the reader must cycle back through the process. This might require the reader to re-identify features in a question or directive or to re-enter the document and to search and locate additional features. Once the reader determines that sufficient information has been matched, the task can be executed by completing the directive.

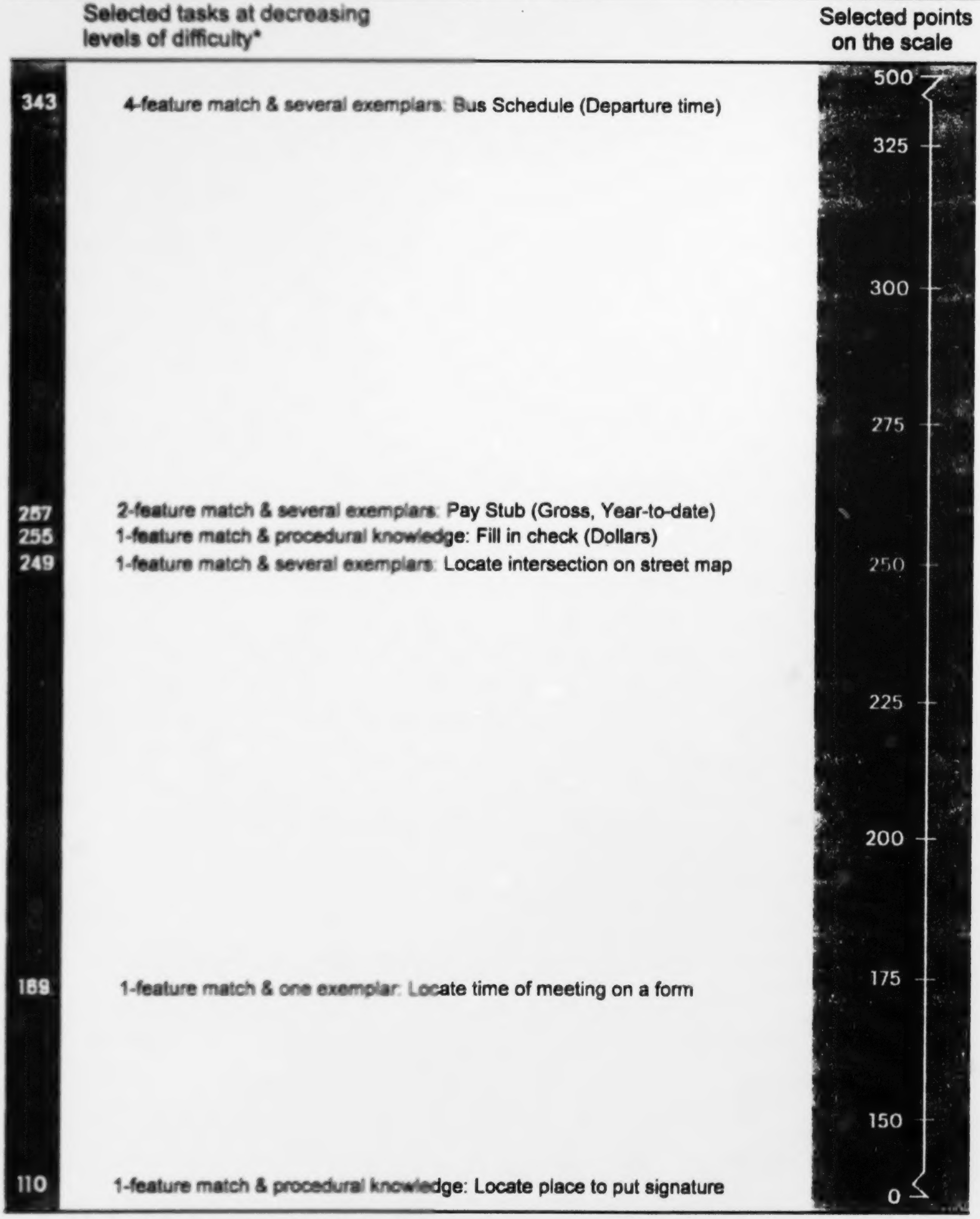
Exhibit 13-4 presents information about task difficulty for the document literacy scale, again based on an item map reported in the young adult literacy assessment (Kirsch and Jungeblut, 1986a). Proficiency in matching document information represents a continuum defined by the number of features that readers must identify in a question or directive. At the simplest end of the continuum, readers match literal information on the basis of a single feature in a document that contains only one exemplar. At the middle and upper ends, the match involves increasing numbers of features. In some cases the matches are literal, while in others the matching is based on varying degrees of correspondence.

Task difficulty increases along with increases in the number of features to be matched, the number of exemplars serving as distractors, and the degree to which information in the question or directive lacks correspondence or identity with the needed information in the document. This aspect of document literacy tasks not only had a dominant role in defining the difficulty of document literacy tasks, but also in defining task difficulty on the prose comprehension scale and, to a lesser extent, on the quantitative scale—that is, matching information in a question or directive with literal or corresponding (synonymous) text information.

13.2.5 Quantitative Literacy Scale

Exhibit 13-5 presents information about task difficulty and population performance, this time on the quantitative literacy scale, based on an original item map from the young adult literacy assessment (Kirsch and Jungeblut, 1986a). Again, it can be seen that 90 percent or more of the total group of young adults surveyed demonstrated proficiency on the tasks extending down from 225 on the quantitative literacy scale. On this scale there were no tasks associated with only an 80 percent success rate for points below 225. Subsequent literacy assessments developed literacy tasks capable of distinguishing using the 80 percent criterion, at the lower end of the scale.

Exhibit 13-4. Selected tasks and corresponding levels of difficulty* defining the document literacy scale: Adults 21 to 25 years old, 1985



*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

Exhibit 13-5. Percentages of adults and selected tasks at or above successive points on the quantitative literacy scale: Adults 21 to 25 years old, 1985

	Selected tasks at decreasing levels of difficulty*	Selected points on the scale	Percent of total
489	Determine amount of interest charges from loan ad	500	
		400	
376	Estimate cost using grocery unit-price labels	375	9.5 (0.9)
371	Calculate & total costs based on item costs from catalogue		
356	Determine tip given percentage of bill		
		350	22.5 (1.4)
340	Plan travel arrangements for meeting using flight schedule		
337	Determine correct change using menu		
		325	37.8 (1.6)
		300	56.0 (1.4)
281	Enter & calculate checkbook balance		
		275	72.2 (1.1)
		250	84.7 (1.0)
233	Total bank deposit entry		
		225	92.4 (0.6)
		0	

*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

The 15 tasks comprising the 1985 quantitative literacy scale appeared to reflect the ability to use mathematical operations such as addition, subtraction, multiplication, or division, either singly or in combination, to solve problems variously embedded in printed material.

Factors associated with task difficulty and performance on the quantitative scale appeared to be the type of arithmetic operation (addition, subtraction, multiplication, and division) required for a correct answer, the number or combination of operations needed, and the extent to which the specification of the operations are embedded in textual material. At the simplest end of the continuum, readers carry out a single, specified operation on numbers that appear in convenient places on the document. At the next level of difficulty, tasks require a single operation, but they also require that the reader enter the appropriate information from the question or directive onto the document before the operation can be completed. At a more difficult level, tasks require either two sequential operations or the application of a single, higher level operation (multiplication or division). At the upper end, the tasks require disembedding the appropriate features of a problem (in the presence of distractors) and carrying out a sequence of operations.

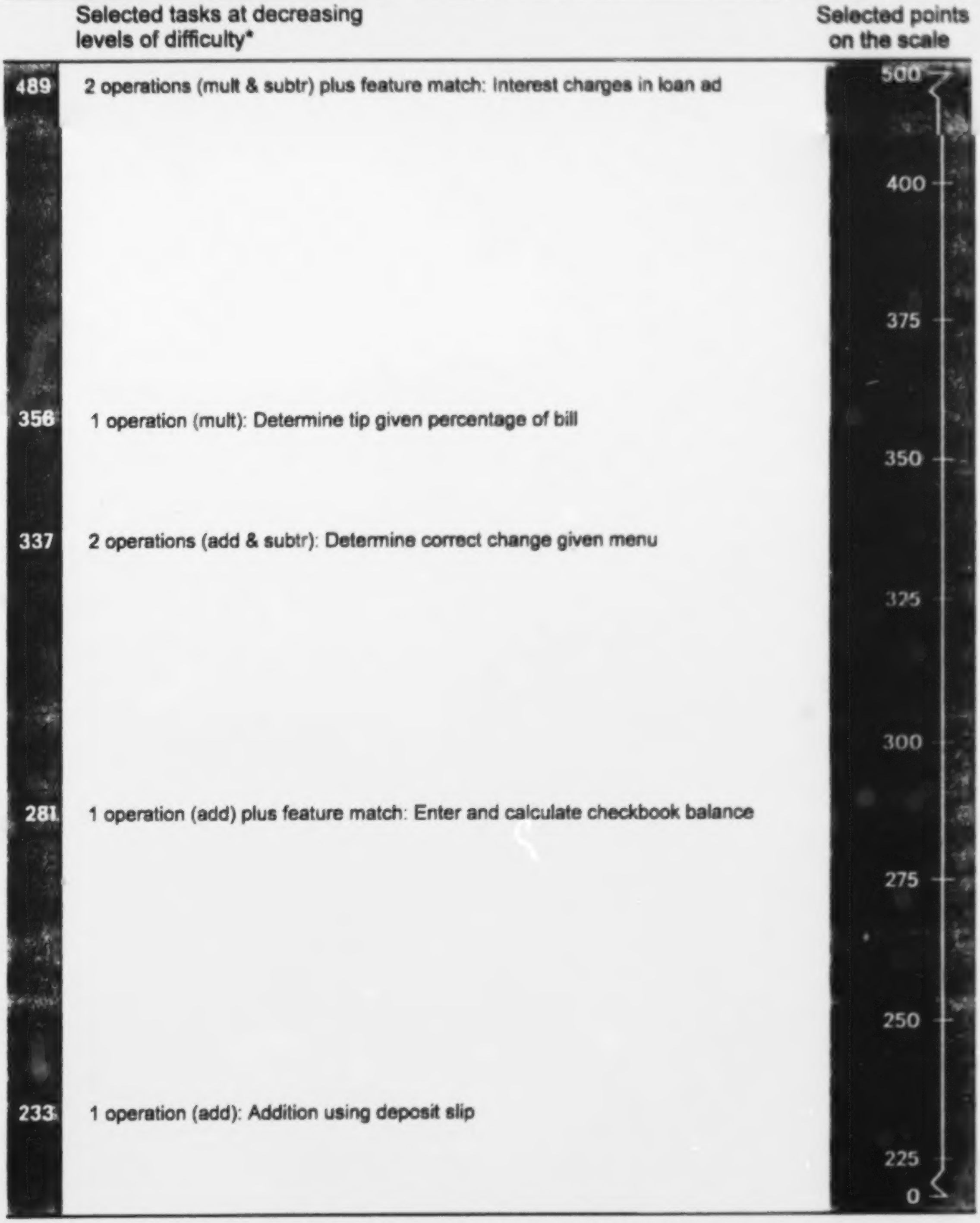
Task difficulty on the quantitative scale was associated not only with the type of operation but also with the number of operations required and the degree to which the problem is embedded in printed material. Exhibit 13-6 presents information about task difficulty for the quantitative literacy scale, based again on an item map reported in the 1985 young adult literacy assessment (Kirsch and Jungeblut, 1986a).

13.3 ENHANCING UNDERSTANDING OF TASK DIFFICULTY

Following publication of the final report from the 1985 young adult literacy assessment, Kirsch and Mosenthal undertook a secondary analysis of the 1985 assessment data in order to extend the understanding of factors associated with document task difficulty—that is, the location of tasks along the document literacy scale (Kirsch and Mosenthal, 1990). They applied an extensive grammar to the 37 different stimulus documents (representing nine categories, e.g., tables, graphs, charts, and forms) as well as to the questions or directives of the 61 specific document literacy tasks developed for the young adult assessment. Their intent was to describe the structure and content of diverse printed materials.

The labor-, knowledge-, and time-intensive grammar devised by Kirsch and Mosenthal relied on semantic-relation categories drawn from other propositional grammars (Mosenthal, 1985). Their application of the grammar to both stimulus materials and associated task directives revealed structural patterns in the document tasks, patterns that encouraged Kirsch and Mosenthal to hypothesize about the variables underlying performance on the literacy tasks. These hypothesized variables related to the structure and complexity of the document or stimulus material, to the nature of the task (i.e., the structural relations between the document and the question or directive), and to the nature of the processes readers engaged in to relate information in the question or directive to information in the document. They

Exhibit 13-6. Selected tasks and corresponding levels of difficulty* defining the quantitative literacy scale: Adults 21 to 25 years old, 1985



	Selected tasks at decreasing levels of difficulty*
489	2 operations (mult & subtr) plus feature match: Interest charges in loan ad
356	1 operation (mult): Determine tip given percentage of bill
337	2 operations (add & subtr): Determine correct change given menu
281	1 operation (add) plus feature match: Enter and calculate checkbook balance
233	1 operation (add): Addition using deposit slip

*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

identified six document variables, four task variables, and three process variables they believed would contribute to difficulty in processing documents. Their initial analyses, using percent correct statistics (rather than response probabilities), identified 12 of the 13 hypothesized variables within their three major categories of variables as significantly influencing the demonstrated difficulty of document literacy tasks for young adults.

Kirsch and Mosenthal further reduced these 12 variables down to a set of eight by omitting those variables with a zero-order correlation of less than .30 with percent correct. Subsequent regression analysis of these eight variables showed that there were five variables that contributed significantly to variance in the percentage correct scores and were consistent (i.e., generalizable) across both racial/ethnic groups and levels of educational attainment (Kirsch and Mosenthal, 1990, Table 2).

Of the final set of five significant variables, two were process variables: degree of correspondence and type of information (Mosenthal and Kirsch, 1991a). *Degree of correspondence* refers to the first stage of document processing, in which a reader must match information given in a question with corresponding information in a document, and varies from easiest (literal or synonymous correspondence) to hardest (correspondence arrived at via high, text-based inference or using special prior knowledge) (Kirsch and Mosenthal, 1990, p. 19). While the *degree of correspondence* variable deals with the correspondence between the information given in the question and the information contained in the document, *type of information* focuses primarily on the requested information. More specifically, type of information refers to how the reader obtains that information, by locating, identifying, generating, or synthesizing requested information based on various "nodes" of a document's information hierarchy. Document processing becomes more difficult as: (a) the reader must generate inferences or use prior knowledge to relate the request to the document, or (b) the reader must relate information across different nodes to arrive at a response (Kirsch and Mosenthal, 1990, p. 20).

Of the final set of five significant variables, two others were task variables (the number of organizing categories and the number of specifics that needed to be processed to complete a task successfully). These variables represent the two types of structural relations between a question or directive and the document and quantify aspects of the amount of information that the reader must process. The *number of organizing categories* consists of the quantity of labels that serve to summarize or synthesize specific data or entries in a document. The *number of specifics* deals with the number of entries or pieces of information the reader must process in order to respond correctly to the task.

Of the final set of five significant variables, the remaining variable was a document variable (the number of specifics), involving the length and complexity of the document itself. The *number of specifics* was a measure of the length and amount of material that was contained in the document. As the number of specifics increased, so did the difficulty of the document.

One additional process variable was notable, though it was not among the final set of five significant variables—plausibility of distractors. This variable refers to the situation where information in a text or document meets some but not all of the conditions required in a question or directive to provide a correct response. Despite the fact that this variable failed to reach significance for the various racial/ethnic and educational attainment groups, it had one of the highest zero-order correlations with percent correct scores among various subgroups of interest.

With the exception of young adults reporting zero to eight years of education, the variance in task difficulty accounted for by the subset of five significant variables ranged from 89 percent (for the total group assessed and for White young adults) to 81 percent (for both black young adults and those who dropped out of high school before earning a diploma). Some 56 percent of the variance in percentage correct scores was accounted for in the group of individuals reporting zero to eight years of schooling. In general, then, the results yielded strong empirical validity evidence for both the evolving theory and for document score interpretation.

This study provided not only a theoretically-based model of performance but also an applied means of predicting task difficulty, along with identified cognitive characteristics for the set of literacy tasks included in the 1985 young adult literacy assessment (Mosenthal and Kirsch, 1991). Although not reported in the literature because of the relatively small number of tasks, the same approach was also applied to both the prose and quantitative scales. Given their enhanced understanding of a set of variables that seem to underlie successful performance on document literacy tasks, Kirsch and Mosenthal used this knowledge to devise specifications for developing new tasks targeted to specific degrees of difficulty along the literacy scales.

The original coding scheme was useful in the design and development of new tasks written for the 1990 survey of the literacy skills of Job Training Partnership Act (JTPA) and Employment Service/Unemployment Insurance (ES/UI) program participants. Additional evidence for the validity of the theory and for score interpretation rests on the success of this task development work.

13.4 THE 1990 SURVEY OF THE LITERACY OF JOB-SEEKERS

The 1990 survey of the literacy skills of job-seekers served by the U.S. Department of Labor (Kirsch and Jungeblut, 1992) capitalized on the results of the 1985 young adult literacy assessment, as well as on the secondary data analyses conducted by Kirsch and Mosenthal. The definition of literacy, the three literacy scales, and the expanded theoretical framework all contributed to the 1990 survey. While the earlier developments provided an important initial step, several of the variables required using the complex and labor intensive grammar. Through a series of revisions and enhancements, a set of variables was identified that eliminated the need for using the grammar and greatly improved the utility of the coding procedures for others interested in the area of literacy. Before discussing the 1990 survey, it will be useful to briefly describe how the current set of variables for prose, document, and quantitative tasks evolved from the original research.

The process variables Kirsch and Mosenthal identified as important through their secondary analysis included *degree of correspondence*, *type of information*, and *plausibility of distractors*. *Plausibility of distractors* was the only variable to remain unchanged throughout this process. A new process variable, *type of match*, was developed by merging the original *type of information* variable with the *degree of correspondence* variable. A third process variable, *type of information*, was also added to indicate the degree of abstractness of the information requested in a question or directive.

Finally, there was some concern that the process variables identified and the associated variance being accounted for might possibly reflect simply the notion of "readability," which has a long history in theoretical and applied research. To address this issue, an estimate of readability was devised from the grammar for use with document stimuli (Mosenthal and Kirsch , 1998), and from the Fry (1977) formula for use with prose stimulus materials.

The following sections evaluate the utility of the current framework as it relates to the creation of new tasks, the understanding of the variables contributing to task difficulty, and the enhancement of score meaning in the 1990 survey of the literacy skills of job-seekers.

13.4.1 Prose Literacy

An important area of literacy is the knowledge and skills needed to understand and use information organized in sentence and paragraph formats. Given the range of text types organized in such formats, the 1990 job-seeker assessment used prose materials that were primarily expository (i.e., materials which describe one or more states or actions) since such materials constitute much of the prose that adults read (Kirsch and Jungeblut, 1986a; Kirsch et al., 1992). In addition, some narrative texts and poetry were included. The prose materials were drawn from newspapers, magazines, books, brochures, and pamphlets, and were reprinted in their entirety, using the typography and layout of the original source. As a result, the materials varied widely in length, in density of information, and in the use of structural or organizational aids, such as section or paragraph headings, italic or bold face type, and bullets.

13.4.1.1 Prose variables

Prose tasks involve the problem of first identifying *given* and *requested* information (Fisher, 1981; Clark and Haviland, 1977; Mosenthal and Kirsch, 1991). *Given* information is that which is known and assumed to be true based on the way a question or directive is stated. *Requested* information in a question or directive is that which is being sought. To illustrate this, consider the question, "In the past five years, how many times has Susan Butcher won the Iditarod Sled Dog Race?" The *given* information in this instance is "In the past five years, Susan Butcher won the Iditarod Sled Dog Race one or more times." The *requested* information of this sentence is "*How many times* did Susan Butcher win?" In processing prose, tasks tend to be easy when the requested information is concrete; tasks tend to be more difficult the more abstract the requested information becomes. Hence, a task whose requested information involves a person or thing (e.g., a *who* or *what* question) tends to be easier to process than a task whose requested information asks for a reason, purpose, or cause (e.g., a *why* question) (Mosenthal, 1998).

Another dimension of prose processing requires readers to match information in a question or directive to corresponding information in a text. This involves the strategies of *locating, cycling, integrating,* and *generating* information. *Locating* tasks require the reader to find information in the text based on conditions or features specified in the question or directive. The match may be literal or synonymous, or the reader may need to make an inference in order to perform successfully. *Cycling* tasks require the reader to locate and match one or more features but also require the reader to engage in a series of feature matches to satisfy conditions given in the question. *Integrating* tasks require the reader to compare or contrast two or more pieces of information from the text. In some cases the information can be found in a single paragraph, while in others it appears in different paragraphs or sections. In the *generating* tasks, readers must produce a written response by processing information from the text and also by making text-based inferences or drawing on their own background knowledge. These processes are represented by the variable *type of match* (Mosenthal, 1998).

A third dimension of prose processing involves *plausibility of distractors,* the situation where information in text meets some but not all the conditions of the answer specified in the question or directive. The more conditions that such distracting information shares with a correct answer and the more closely it is positioned to the correct answer, the more difficult the processing becomes (Mosenthal, 1996).

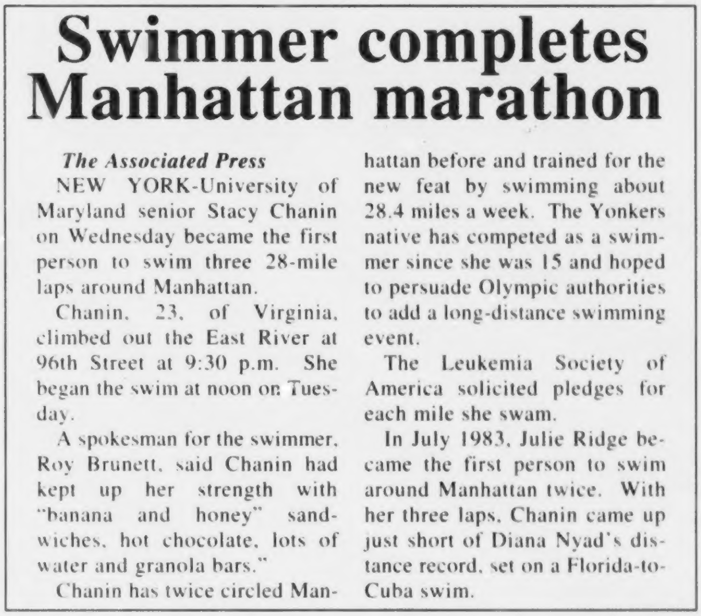
In addition to the three process variables, Kirsch and Mosenthal considered a fourth variable—readability—representing complexity of prose materials (Fry, 1977). It was included not only to determine the extent to which it accounted for task difficulty, but also to provide another descriptor commonly found in the research literature on prose processing (Mosenthal, 1998).

Kirsch and Mosenthal evaluated each of the 45 prose tasks included in the 1990 job-seeker assessment in terms of these four variables. They devised a coding scheme and applied it to each of the prose literacy tasks. The variables *type of information* and *plausibility of distractors* range from one (easiest) to five (most difficult). The coding scheme they applied to *type of match* was additive and ranged from one to a possible 20, although the type of match variable for actual tasks in the 1990 survey ranged from only one to eight. Based on this scheme, several example tasks are presented in the next section that highlight the range of task complexity required for successful performance along this dimension of literacy.

13.4.1.2 Examples of prose literacy tasks

One of the easiest prose tasks (RP80 difficulty value of 210) involved a short newspaper article about a marathon swimmer (Exhibit 13-7). This text reflects an eighth-grade Fry readability level. The directive asks the reader to "underline the sentence that tells what Ms. Chanin ate during the swim." To complete this directive, readers have to recognize that the requested information is a thing (i.e., food). This prose task received a code of 1 for the *type of information* process variable. In identifying the requested information, readers must make a synonymous match between "ate" in the directive and "banana and honey sandwiches, hot chocolate, lots of water, and granola bars" in the text. This task received a code of 1 for the *type of match* process variable. Note that, since there is no other mention of food in the text, there are no plausible distractors for requested information. This task received a code of 1 for the *plausibility of distractors* process variable.

Exhibit 13-7. Example of text for relatively easy prose literacy tasks



Swimmer completes Manhattan marathon

The Associated Press

NEW YORK-University of Maryland senior Stacy Chanin on Wednesday became the first person to swim three 28-mile laps around Manhattan.

Chanin, 23, of Virginia, climbed out the East River at 96th Street at 9:30 p.m. She began the swim at noon on Tuesday.

A spokesman for the swimmer, Roy Brunett, said Chanin had kept up her strength with "banana and honey" sandwiches, hot chocolate, lots of water and granola bars."

Chanin has twice circled Manhattan before and trained for the new feat by swimming about 28.4 miles a week. The Yonkers native has competed as a swimmer since she was 15 and hoped to persuade Olympic authorities to add a long-distance swimming event.

The Leukemia Society of America solicited pledges for each mile she swam.

In July 1983, Julie Ridge became the first person to swim around Manhattan twice. With her three laps, Chanin came up just short of Diana Nyad's distance record, set on a Florida-to-Cuba swim.

Find the article "Swimmer completes Manhattan marathon" on page 2 or the newspaper provided and answer the following questions.

11. Underline the sentence that tells what Ms. Chanin ate during the swim.
12. At what age did Chanin begin swimming competitively? _______

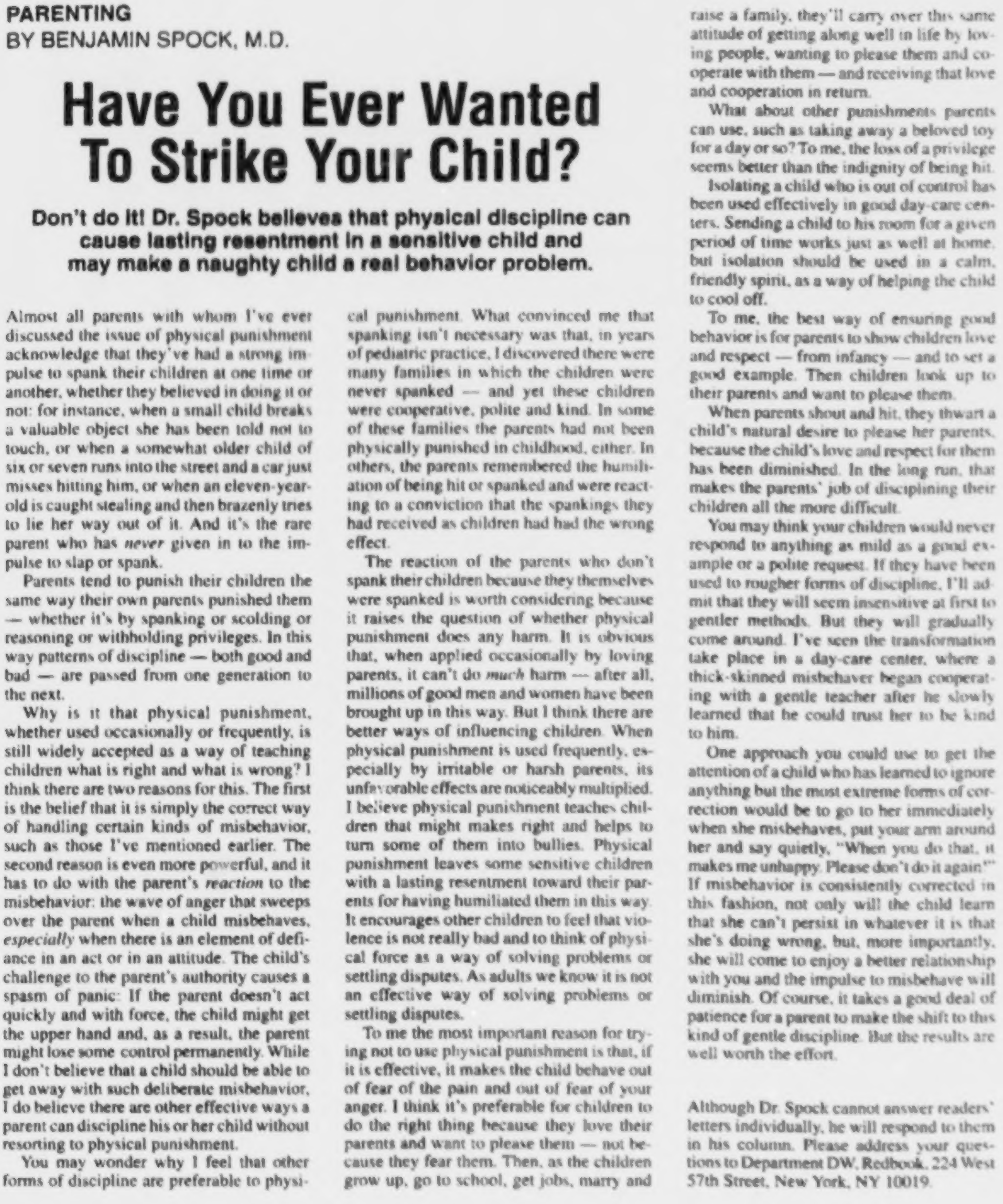
A second task involving this text has an RP80 scale value of 250. This task includes the question, "At what age did Chanin begin swimming competitively?" To answer this question, readers must recognize that the requested information is an amount (i.e., age). This task received a code of 2 for the *type of information* process variable (because an amount in this task is more abstract than a thing in the previous task). To identify the requested information, readers have to make a synonymous match. Having made this match, readers then must find the answer "15" in the sentence "The Yonkers native has competed as a swimmer since she was 15 and Y." This task received a 3 for the *type of match* process variable (because a cycle of matching was required). What also makes this task somewhat difficult is the fact that there is a distractor for the requested information that appears earlier in the text; this distractor includes the information that the swimmer's current age is 23. This task received a 4 for the *plausibility of distractors* process variable (because distractors appear for both given and requested information, but not in the same paragraph as the answer).

A task receiving an RP80 scale value of 247 involved a rather lengthy article on parenting written by Dr. Spock (Exhibit 13-8). While this article only represented an eighth-grade Fry readability level, it consists of a relatively long passage without any organizational aids (a challenge not measured by the Fry approach). One question asked the reader to identify one alternative to the use of physical punishment. The reader could match the phrase "alternative to the use of physical punishment" to the phrase "other punishments parents can use." The text then lists a number of alternatives recommended by Dr. Spock. *Type of match* received a score of 2 (because cycling was required) and *plausibility of distractors* also received a score of 2 (because similar information appears somewhere in the text, but not nearby), while *type of information* received as score of 3 (because an alternative is more abstract than a thing or an amount).

A somewhat more difficult task based on the same text (RP80 score of 283) requires the reader to "list the two reasons given by the author why physical punishment is still widely accepted as a way to teach children right and wrong." This task can be answered by locating the place in the text that begins, "I think there are two reasons for this. The first is .... The second reason is ...." *Type of match* received a score of 3 (an extra point was added to a 2 for cycling by the need to identify the antecedent of the pronoun "this"). *Plausibility of distractors* received a score of 2 (because similar information appears somewhere in the text, but not nearby). *Type of information* received an abstractness score of 4 (because a reason is more abstract than an alternative).

The most difficult task involving this text (an RP80 score of 311) directs the reader to "identify and list two reasons that Dr. Spock offers for not using physical punishment." While numerous statements throughout the article help satisfy the directive, much of the text deals with related concerns rather than direct summary statements. As a result, the reasons for not using physical punishment are embedded throughout the text and are not literally stated following a semantic cue such as "Two good reasons for not using physical punishment are Y." This task was coded 3 for *type of match* (because it requires synthesizing features across the document). In addition, distracting information is more closely tied to words or phrases containing the necessary information for responding correctly. This task was coded 3

Exhibit 13-8. Example of text for moderately difficult prose literacy tasks



PARENTING
BY BENJAMIN SPOCK, M.D.

Have You Ever Wanted To Strike Your Child?

Don't do it! Dr. Spock believes that physical discipline can cause lasting resentment in a sensitive child and may make a naughty child a real behavior problem.

Almost all parents with whom I've ever discussed the issue of physical punishment acknowledge that they've had a strong impulse to spank their children at one time or another, whether they believed in doing it or not: for instance, when a small child breaks a valuable object she has been told not to touch, or when a somewhat older child of six or seven runs into the street and a car just misses hitting him, or when an eleven-year-old is caught stealing and then brazenly tries to lie her way out of it. And it's the rare parent who has *never* given in to the impulse to slap or spank.

Parents tend to punish their children the same way their own parents punished them — whether it's by spanking or scolding or reasoning or withholding privileges. In this way patterns of discipline — both good and bad — are passed from one generation to the next.

Why is it that physical punishment, whether used occasionally or frequently, is still widely accepted as a way of teaching children what is right and what is wrong? I think there are two reasons for this. The first is the belief that it is simply the correct way of handling certain kinds of misbehavior, such as those I've mentioned earlier. The second reason is even more powerful, and it has to do with the parent's *reaction* to the misbehavior: the wave of anger that sweeps over the parent when a child misbehaves, *especially* when there is an element of defiance in an act or in an attitude. The child's challenge to the parent's authority causes a spasm of panic: If the parent doesn't act quickly and with force, the child might get the upper hand and, as a result, the parent might lose some control permanently. While I don't believe that a child should be able to get away with such deliberate misbehavior, I do believe there are other effective ways a parent can discipline his or her child without resorting to physical punishment.

You may wonder why I feel that other forms of discipline are preferable to physical punishment. What convinced me that spanking isn't necessary was that, in years of pediatric practice, I discovered there were many families in which the children were never spanked — and yet these children were cooperative, polite and kind. In some of these families the parents had not been physically punished in childhood, either. In others, the parents remembered the humiliation of being hit or spanked and were reacting to a conviction that the spankings they had received as children had had the wrong effect.

The reaction of the parents who don't spank their children because they themselves were spanked is worth considering because it raises the question of whether physical punishment does any harm. It is obvious that, when applied occasionally by loving parents, it can't do *much* harm — after all, millions of good men and women have been brought up in this way. But I think there are better ways of influencing children. When physical punishment is used frequently, especially by irritable or harsh parents, its unfavorable effects are noticeably multiplied. I believe physical punishment teaches children that might makes right and helps to turn some of them into bullies. Physical punishment leaves some sensitive children with a lasting resentment toward their parents for having humiliated them in this way. It encourages other children to feel that violence is not really bad and to think of physical force as a way of solving problems or settling disputes. As adults we know it is not an effective way of solving problems or settling disputes.

To me the most important reason for trying not to use physical punishment is that, if it is effective, it makes the child behave out of fear of the pain and out of fear of your anger. I think it's preferable for children to do the right thing because they love their parents and want to please them — not because they fear them. Then, as the children grow up, go to school, get jobs, marry and raise a family, they'll carry over this same attitude of getting along well in life by loving people, wanting to please them and cooperate with them — and receiving that love and cooperation in return.

What about other punishments parents can use, such as taking away a beloved toy for a day or so? To me, the loss of a privilege seems better than the indignity of being hit.

Isolating a child who is out of control has been used effectively in good day-care centers. Sending a child to his room for a given period of time works just as well at home, but isolation should be used in a calm, friendly spirit, as a way of helping the child to cool off.

To me, the best way of ensuring good behavior is for parents to show children love and respect — from infancy — and to set a good example. Then children look up to their parents and want to please them.

When parents shout and hit, they thwart a child's natural desire to please her parents, because the child's love and respect for them has been diminished. In the long run, that makes the parents' job of disciplining their children all the more difficult.

You may think your children would never respond to anything as mild as a good example or a polite request. If they have been used to rougher forms of discipline, I'll admit that they will seem insensitive at first to gentler methods. But they will gradually come around. I've seen the transformation take place in a day-care center, where a thick-skinned misbehaver began cooperating with a gentle teacher after he slowly learned that he could trust her to be kind to him.

One approach you could use to get the attention of a child who has learned to ignore anything but the most extreme forms of correction would be to go to her immediately when she misbehaves, put your arm around her and say quietly, "When you do that, it makes me unhappy. Please don't do it again!" If misbehavior is consistently corrected in this fashion, not only will the child learn that she can't persist in whatever it is that she's doing wrong, but, more importantly, she will come to enjoy a better relationship with you and the impulse to misbehave will diminish. Of course, it takes a good deal of patience for a parent to make the shift to this kind of gentle discipline. But the results are well worth the effort.

Although Dr. Spock cannot answer readers' letters individually, he will respond to them in his column. Please address your questions to Department DW, Redbook, 224 West 57th Street, New York, NY 10019.

Benjamin Spock: "Have You Ever Wanted to Strike Your Child?" Reprinted from *Redbook* by permission of the publisher.

BEST COPY AVAILABLE

for *plausibility of distractors*. As with the preceding task, *type of information* received an abstractness score of 4 (because the task requests a reason, rather than an alternative, an amount, or a thing).

A more difficult task (RP80 score value 346) directs the reader to identify and list two similarities between the new and old ways American Express handles charge card receipts (Exhibit 13-9). This piece of text received a Fry readability score of 8 (eighth grade), indicating that this should not be a difficult text for most adults. The correct response requires the reader to compare and contrast several pieces of information to determine these similarities. Because integrating information through comparison and contrast is relatively demanding, this task was coded 4 for *type of match*. When the information requested is a "similarity," as in this task, *type of information* is scored as a 4. *Plausibility of distractors* received a score of 2 (because similar information appears somewhere in the text, but not in the same paragraph as the answer).

Exhibit 13-9. Example of text for difficult prose literacy tasks

American Express' Way of Handling the Flood of Charge Card Receipts

How the new way stacks up against the old way

The New Way:
1 Image processing camera converts receipts to electronic digital image and paper receipts are discarded. 2 Digital image is scanned for account and invoice numbers by optical character (99% accuracy). In the future, computers will also read handwritten charge amounts. 3 Charge amounts are entered by computer operator from image displayed on computer screen. 4 Images are sorted electronically. 5 Bills, with images of receipts, are printed by laser and mailed to cardholders. 6 Images of receipts are stored permanently on optical discs.

The Old Way:
1 Paper receipts are microfilmed for 2 permanent storage, then 3 scanned for account and invoice number by optical character reader (82% accuracy). 4 Charge amounts are entered by computer operator from receipts. 5 A code containing all the information is printed on the receipts. 6 Paper receipts are sorted. 7 Bills are generated by mainframe computer. 8 Receipts and bills are joined and mailed.

One of the most difficult prose tasks in the 1990 survey (RP80 score of 441) required the reader to identify two differences in the new and old ways of handling the American Express receipts. The correct response requires the reader to compare and contrast several pieces of information to determine these differences. Because identifying differences through comparison and contrast is more demanding than identifying similarities, this task was coded 7 for *type of match*. When the information requested is a "difference," as in this task, *type of information* is scored as a 5. This task was also coded as a 5 for *plausibility of distractors*.

The tasks shown above provide examples of how the process variables were assigned numeric values to capture the extent of the various information-processing demands that such prose literacy tasks place on readers.

13.4.1.3 Coding the 1985 and 1990 prose literacy tasks

The preceding section provided examples to illustrate how type of match, type of information, and plausibility of distractors were coded for selected prose literacy tasks from the 1990 literacy survey of job-seekers served by the Department of Labor. In this section, the coding rules are formulated in more general terms. A number of criteria must be taken into account when measuring the four variables associated with task difficulty on the prose scale.

13.4.1.3.1 Type of Match

Description. Type of match refers to the processes used to relate requested information to the corresponding information in a prose text, and to the process of entering a response. Four basic types of match can be distinguished: *locate*, *cycle*, *integrate*, and *generate* matches.

Locate tasks require users to match one or more features in a question to one or more features in the text (Kirsch and Mosenthal, 1992a; Mosenthal, 1998). Based on this match, the answer is located in the appropriate paragraph or sentence of a prose text.

Cycle tasks require users to perform an iterative series of locate matches and may involve the selection of several pieces of information that meet a criterion. With prose texts, cycle tasks are made difficult depending upon whether they are performed within a paragraph or between paragraphs. Cycle tasks are further made difficult depending upon whether the cycles are independent of one another or a sequence in which each answer is used to identify the next part of the locating cycle.

Integrate tasks require users to compare or contrast information that has been located in two or more different locate matches or in one or more cycle matches (Kirsch and Mosenthal, 1992; Mosenthal, 1998). In general, integrate tasks which require readers to compare information are easier than those that require readers to contrast information.

Generate tasks require readers to use prior knowledge (often representing a specific type of content knowledge) to match information in a question or directive to corresponding information in a prose text (Mosenthal and Kirsch, 1993a; Mosenthal, 1998). Moreover, generate tasks may require readers to use specialized knowledge to select from among a set of plausibly correct responses the answer which best meets the conditions stated in a question or directive. Without the benefit of such knowledge, users often must guess or ask some expert to complete the match.

Scoring rules. The scoring of type of match is basically determined by the nature of the match, and to a lesser extent by the number of phrases or features in the request, the number of responses requested, whether prior knowledge is needed for the required inferences, and how a choice among several possible answers should be selected. The basic idea, however, is that locate matches are easier than cycle matches, cycle matches are easier than integrate matches, and integrate matches are easier than generate matches.

Exhibit 13-10. Basic scoring rules for type of match: Nature of the task

Rule	Score
When the task is to *locate* the information in the prose text or document that corresponds to the features requested.	1
When the task is to *cycle* (that is, perform an iterative series of locate matches) to find the information that corresponds to the features requested. Add 1 point if the answer is located in more than one paragraph.	2
When the task is to *integrate* information located in a prose text by comparing, or for prose text, when the task is to infer a condition based on a synthesis of features found in the same paragraph of text.	3
When the task is to integrate information located in a prose text by contrasting, or for prose text, when the task is to infer a condition based on a synthesis of features found in more than one paragraph of text.	4
When the task is to *generate* new information (that is, to use prior knowledge to match information requested with that in the prose text).	5

Sometimes matching is made more difficult as the number of *phrases or features* in the directions required to locate an answer increases (Mosenthal, 1998). Matches that require the identification of only a single phrase or feature are, on average, easier than matches that require the identification of two phrases or features. Given the basic score based on the nature of the match, additional points can be added based on the number of phrases or features in the directions.

Exhibit 13-11. Additional scoring rule for type of match: Number of phrases or features in request

Rule	Add
When the request for information consists of one independent clause and one dependent clause	1
When the request for information consists of one independent clause and two dependent clauses	2
When the request for information consists of one independent clause and three or more dependent clauses	3

Sometimes matching is also made more difficult as the number of *responses* readers must supply increases and as the specificity of this number decreases. Requests of readers to list only one answer are easier than requests to list two or three answers; requests of readers to list two or three answers are easier than requests for four answers. Requests that do not specify the number of responses explicitly are harder than those that do specify the number of multiple responses required. Given the preliminary score based on the nature of the match and the number of phrases or features, additional points can be added based on the number of responses requested.

Exhibit 13-12. Additional scoring rule for type of match: Number of responses requested

Rule	Add
When readers are requested to list two or three answers	1
When readers are requested to list four or five answers $ add 1 point if the request does not specify exactly how many	2
when readers are requested to list six or more answers $ add 1 point if the request does not specify exactly how many	3

Matching can be further made difficult to the extent that readers have to *make inferences* to match information in the question to information in the document or text (Mosenthal and Kirsch, 1993b). Questions may require either low text-based inferences (which consist of inferences to be made within the context of information provided in a text), or high text-based inferences (which consist of inferences requiring some combination of knowledge of the text and specialized prior knowledge). Low text-based inferences are easier to make than high text-based inferences. Given the preliminary score based on the nature of the match, the number of phrases or features, and the number of responses requested, additional points can be added based on the kind of inference needed to answer.

Exhibit 13-13. Additional scoring rule for type of match: Inferences needed

Rule	Add
When the text alone provides sufficient information to make an inference needed to match the request with the information in the text or document (a low text-based inference)	1
When prior knowledge as well as the text is needed to make the inference needed to match the request with the information in the text or document (a high text-based inference)	3

Sometimes matching is made more difficult when all possibilities match the request and readers have to choose which one of several possible answers best completes a requested *information frame* (Mosenthal and Kirsch, 1991). In these cases, the match is more difficult when this choice requires using the text to infer why one of several possible answers best completes a requested information frame, when this choice requires identifying conditional information which renders one of the possible answers more consistent with the conditions requested than others, or when this choice requires readers to relate a pronoun to its antecedent before an answer can be provided. Given the preliminary score based on the nature of the match, the number of phrases or features, the number of responses requested, and the kind of inference needed, additional points can be added based on how the reader must complete an information frame.

Exhibit 13-14. Additional scoring rule for type of match: Completing an information frame

Rule	Add
When the choice among candidate answers requires a low, text-based inference (the text alone provides sufficient information to make the inference), the identification of a condition, the identification of a pronoun antecedent, or a restatement of a type of information	2
When the choice among candidate answers requires a high, text-based inference (prior knowledge as well as the text is needed to make the inference)	4

These scoring rules are additive (Meyer, Marsiske, and Willis, 1993). A prose literacy task, for example, might have a basic score of 2 because it is a cycle task, but have additional points added because the cycling occurs between paragraphs (add 1), involves a two-clause question (add 1), needs a low text-based inference (add 1), for which the answer should consist of two responses (add 1), but whose actual number is not explicitly specified (add 1). A prose assessment task with these features would have a total type-of-match score of 7.

The actual prose-based tasks used in this assessment scored from 1 to 8 on type of match. Eight was not a ceiling set in advance. Rather, these upper bounds reflects the range of difficulty combinations which commonly characterize tasks found in society and the workplace. While more difficult tasks could be conceived in designing assessments (for example, a four-phrase contrast task requiring high text-based inferencing and six uncued responses), such tasks would be so difficult that they would bear little resemblance to ordinary usage of prose texts.

13.4.1.3.2 Type of Information

Description. Type of information refers to the degree of concreteness of the objects described in a prose text or document (Mosenthal and Kirsch, 1991b; Mosenthal, 1998). More abstract objects are harder for readers to identify and understand.

Scoring rules. Assessment tasks can be rated in terms of the concreteness of the information requested. Most concrete were tasks requesting information about persons, groups, animals, locations, and things. Somewhat less concrete were questions requiring the identification of amounts, times, attributes, types, actions, locations, and parts. Somewhat abstract were questions requesting information about manner, goals, purposes, alternatives, conditions, pronoun references, and predicate adjectives. Abstract tasks requested the identification of causes, effects, reasons, evidence, similarities, and explanations. Finally, very abstract tasks requested the identification of equivalences, differences, themes, or patterns.

Exhibit 13-15. Scoring rules for type of information

Rule	Score
When the information requested refers to a person, group, animal, place, or thing (most concrete)	1
When the information requested refers to an amount, time, attribute, type, action, location, or part	2
When the information requested refers to a manner, goal, purpose, alternative, attempt, condition, pronominal reference, predicate adjective, sequence, assertion, problem, solution, role, or process	3
When the information requested refers to a cause, effect, reason, result, evidence, similarity, explanation, opinion, or procedure	4
When the information requested refers to an equivalence, difference, theme, pattern, definition, or advantage (most abstract)	5

13.4.1.3.3 Plausibility of Distractors

Description. Distractors are elements of a question's given or requested information that appear in the prose text that, when identified lack an essential feature and do not qualify as correct. Given information is provided by the directions for a task and is used to search for the requested information (that being sought). Unless the possible but incorrect answers are plausible (by sharing some of the features of the correct answer), they do not function as distractors (Mosenthal, 1996, 1998).

Scoring rules. Tasks are easiest when the prose text contains no information related to the conditions set forth in the question other than the answer. Tasks become slightly more difficult when a distractor for either given or requested information (but not both) appears, but does not occur very close to the correct answer. Tasks become more difficult when plausible distractors for both given and requested information appear, but are not both located next to the correct information. This occurs in prose texts when they appear in different paragraphs, one of which may be in the paragraph in which the answer occurs.

Tasks become still more difficult when plausible distractors for both given and requested information appear in the same place but are not located near the correct information. This occurs in prose texts when they both appear in the same paragraph but one other than the paragraph in which the answer appears. Tasks involving prose texts also reach this level of difficulty when negatives become involved—the plausible distractors represent the opposite condition of what is established in the question or directive, and these distractors appear in a paragraph other than the one in which the answer occurs.

Tasks are most difficult when plausible distractors for both given and requested information appear in the same place, or appear with negative conditions, and are located near the correct information. This occurs in prose texts when they both appear in the same paragraph as the one in which the answer occurs, or when the distractors represent the opposite condition of what is established by the task, and they appear in the same paragraph as the answer.

Exhibit 13-16. Scoring rules for plausibility of distractors

Rules for prose texts	Score
When no information related to the conditions requested appears, other than the answer (no plausible distractors)	1
When information similar to either given or requested information appears somewhere in the text but not near the answer, or inferences invited by information in the paragraph containing the answer bear a resemblance to the answer	2
When distractors for both given and requested information appear in different paragraphs, though one could occur in the paragraph containing the answer	3
When distractors for both given and requested information, or when plausible distractors represent the opposite condition of what is requested, appear in the same paragraph, but one other than the paragraph containing the answer	4
When distractors for both given and requested information, or when plausible distractors represent the opposite condition of what is requested, appear in the same paragraph as the answer	5

13.4.1.3.4 Readability

Description. Prose texts vary in the length of sentences, number of syllables in the words used, and the complexity of the syntax, while documents vary in their complexity, depending on their organization, number of elements, and number of labels. Literacy tasks may be easier to process when the structure of the document or prose text containing the needed information is less complex. The measurement of the complexity of prose texts derives from Fry's research on readability (Fry, 1975, 1977, 1981).

Scoring rules for prose text. Readability of prose is based on the average number of syllables per 100 words and the average number of sentences per 100 words. These two continuous variables are then used as coordinates in Fry's (1977) readability grade level graph, which portrays a nonlinear relationship between the two and the resulting readability level. In general, however, the more syllables per word and the more words per sentence, the higher the associated grade level of the text. The grade levels of the texts used in the National Adult Literacy Survey ranged from fourth to fifteenth.

13.4.1.4 Codes for all 1985 and 1990 prose literacy tasks

The preceding sections described in detail the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the prose scale. These rules were applied to all prose literacy tasks in the 1990 survey and in the 1985 young adult literacy assessment. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-2.

Table 13-2. List of prose literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: 1990 survey of the literacy of job-seekers

Identifier	Task Description	Scaled RP80	IRT parameters			Readability	Type of match	Distractor Plausibility	Information type
			a	b	c				
A111301	Toyota, Acura, Nissan	189	0.868	-2.488	0.000	8	1	1	1
AB21101	Swimmer: Underline sentence telling what Ms. Chanin ate	208	1.125	-1.901	0.000	8	1	1	1
A120501	Blood donor pamphlet	216	0.945	-1.896	0.000	7	1	1	2
A130601	Summons for jury service	237	1.213	-1.295	0.000	7	3	2	2
A120301	Blood donor pamphlet	245	0.956	-1.322	0.000	7	1	2	3
A100201	PHP subscriber letter	249	1.005	-1.195	0.000	10	3	1	3
A111401	Toyota, Acura, Nissan	250	1.144	-1.088	0.000	8	3	2	4
A121401	Dr. Spock column: alterntv to phys punish	251	1.035	-1.146	0.000	8	2	2	3
AB21201	Swimmer: Age Ms. Chanin began to swim competitively	250	1.070	-1.125	0.000	8	3	4	2
A131001	Shadows Columbus saw	280	1.578	-0.312	0.000	9	3	1	2
AB80801	Illegal questions	265	1.141	-0.788	0.000	6	3	2	2
AB41001	Declaration: Describe what poem is about	263	0.622	-1.433	0.000	4	3	1	3
AB81101	New methods for capital gains	277	1.025	-0.638	0.000	7	4	1	3
AB71001	Instruction to return appliance: Indicate best note	275	1.378	-0.306	0.266	5	3	2	3
AB90501	Questions for new jurors	281	1.118	-0.493	0.000	6	4	2	1
AB90701	Financial security tips	262	1.563	-0.667	0.000	8	3	2	4
A130901	Shadows Columbus saw	282	1.633	-0.255	0.000	9	3	4	1
AB60201	Make out check: Write letter explaining bill error	280	1.241	-0.440	0.000	7	3	2	4
AB90601	Financial security tips	299	1.295	-0.050	0.000	8	2	2	4
A121201	Dr. Spock column: why phys punish acceptd	285	1.167	-0.390	0.000	8	3	2	4
AB70401	Almanac vitamins: List correct info from almanac	289	0.706	-0.765	0.000	7	3	4	1
A100301	PHP subscriber letter	294	0.853	-0.479	0.000	10	4	3	2
A130701	Shadows Columbus saw	298	1.070	-0.203	0.000	9	3	2	3
A130801	Shadows Columbus saw	303	0.515	-0.929	0.000	9	3	2	2
AB60601	Economic index: Underline sent. Explaining action	305	0.809	-0.320	0.000	10	3	2	4
A121301	Dr. Spock column: 2 cons against phys punish	312	0.836	-0.139	0.000	8	3	3	4
AB90401	Questions for new jurors	300	1.230	-0.072	0.000	6	4	2	3
AB80901	Illegal questions	316	0.905	0.003	0.000	6	4	3	3
A111101	Toyota, Acura, Nissan	319	0.772	-0.084	0.000	8	4	3	2
AB40901	Korean Jet: Give argument made in article	329	0.826	0.166	0.000	10	4	4	4
A131101	Shadows Columbus saw	332	0.849	0.258	0.000	9	5	4	1
AB90801	Financial security tips	331	0.851	0.236	0.000	8	5	5	2
AB30601	Technology: Orally explain info from article	333	0.915	0.347	0.000	8	4	4	4
AB50201	Panel: Determine surprising future headline	343	1.161	0.861	0.196	13	4	4	4
A101101	AmerExp: 2 Similarities in Handling Receipts	346	0.763	0.416	0.000	8	4	2	4
AB71101	Explain difference between 2 types of benefits	348	0.783	0.482	0.000	9	6	2	5
AB81301	New methods for capital gains	355	0.803	0.652	0.000	7	5	5	3
A120401	Blood donor pamphlet	358	0.458	-0.056	0.000	7	4	5	2
AB31201	Dickinson: Describe what is expessed in poem	363	0.725	0.691	0.000	6	6	2	4
AB30501	Technology: Underline sentence explaining action	371	0.591	0.593	0.000	8	6	4	4
AB81201	New methods for capital gains	384	0.295	-0.546	0.000	7	2	4	2
A111201	Toyota, Acura, Nissan	404	0.578	1.192	0.000	8	8	4	5
A101201	AmExp: 2 Diffs in Handling Receipts	441	0.630	2.034	0.000	8	7	5	5
AB50101	Panel: Find information from article	469	0.466	2.112	0.000	13	6	5	4

13.4.1.5 Validity evidence for the prose scale

One important piece of validation evidence can be obtained from analyses of the tasks used in the 1990 job-seeker survey. As indicated earlier, the prose area was not especially well defined in the 1985 young adult literacy assessment—the scale contained only 15 tasks. As a result, about 30 new prose tasks were developed for the 1990 survey. Each of the new tasks was coded using the revised theoretical framework described in the preceding section. These variables were then used in regression analyses designed to predict the scale values of each task. Table 13-3 shows the results of these analyses for both the new tasks as well as the entire 1990 pool of prose literacy tasks (1985 and 1990 tasks). The numbers in the tables represent the raw beta coefficients for each of the variables included in the regression analyses along with standard errors and probabilities. Overall, the three process variables were significant for both the new 1990 tasks and for the complete set of prose tasks. Although not shown here, readability was significant if entered into the regression by itself and accounted for about 20 percent of the variance in predicting task scale values. However, when combined with the three process variables, it did not increase the explained variance.

Table 13-3. Unstandardized regression coefficients and standard errors predicting RP80 task difficulties on the basis of four structure and process variables: 30 new prose literacy tasks and 44 total prose literacy tasks from the 1990 survey of job-seekers

	New Tasks			All Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Readability	-.04	3.97	.93	2.14	2.35	.37
Process Variables						
Type of match	12.08	5.94	.05	17.68	4.48	.00
Plausibility of distractors	28.53	5.79	.00	20.98	4.56	.00
Type of information	14.35	5.16	.01	12.19	4.28	.01
Variance accounted for:						
R^2		81%			82%	
Adjusted R^2		78%			80%	
Degrees of freedom		25			39	

Using the expanded theoretical framework for task development on the prose scale appears to have been successful. The amount of variance accounted for in the new tasks (81 percent), as well as in the complete set of tasks (82 percent), compares favorably with the research results reported by Kirsch and Mosenthal (1990)—in the range from 81 to 89 percent for document tasks. These results also suggest that readability is less important than the process variables in explaining task difficulty.

13.4.2 Document Literacy

An additional aspect of being literate in today's society is having the knowledge and skills needed to process documents, or information organized in matrix structures (i.e., in rows and columns). Included among documents are such things as tables, signs, indexes, lists, coupons, schedules, charts, graphs, maps, and forms. In contrast to prose, which tends to be the predominant form of literacy in schools, documents tend to be the principal form of literacy in non-school settings (Guthrie, Seifert, and Kirsch, 1986). Documents serve many important functions in our daily lives (Bassett, Goodman, and Fosegan, 1981; Burch and Grudnitski, 1986). They enable people to perform important actions (e.g., applying for benefits, opening a charge account), make informed decisions (e.g., using a table of benefits to determine whether certain medical costs are covered), and record actions (e.g., completing a deposit slip or bill of sale, receiving a ticket for speeding).

13.4.2.1 Document variables

Document literacy tasks require readers to locate and match information in a question or directive to corresponding information in complex arrays, and to use this information in appropriate ways. For example, procedural knowledge may be needed to transfer information from one source or document to another, as is necessary in completing applications or order forms. This matching again involves the strategies of *locating, cycling, integrating,* and *generating* information; these strategies are again represented by the variable *type of match* (Mosenthal, 1996). As with prose literacy tasks, success in processing documents also appears to depend on the ability to identify different *types of information.* Similarly, both prose and document tasks are made more difficult through the presence of *plausible distractors.*

In addition to these three process variables, it was deemed important to provide an index of the readability of document stimuli. Since no such index is readily available for estimating document complexity, Mosenthal and Kirsch (1998) developed a means of estimating this complexity based on the grammar used in the earlier research.

The basic structural unit of documents is "simple lists" (Mosenthal and Kirsch, 1989a, 1998). Such lists consist of a series of exemplars or items that belong to a common class of elements that, in most instances, are organized in terms of a more generic category or label. The documents used in the assessments reflect the ways in which a number of simple lists are organized to present more interdependent sets of information. These lists have been described as combined, intersecting, and nested (Kirsch and Mosenthal, 1989, 1990b; Mosenthal and Kirsch, 1989b, 1998). The document readability variable ranges from 1 to 11. Included in this score is a number reflecting the type of document structure, the number of labels, and the numbers of items.

13.4.2.2 Examples of document literacy tasks

One of the easier tasks on the document scale (RP80 score of 198) required the reader to look at a theater trip notice containing information about two plays (Exhibit 13-17). This document received a low structural complexity score of 2. The reader is directed to circle the cost for a ticket and bus trip to see *On the Town*. Although the reader simply locates the line in the notice labeled "price" and circles the dollar amount associated with *On the Town*, the cost given in the document for *Sleuth* serves as a plausible distractor. This task received a code of 1 for *type of match*, and codes of 2 for *plausibility of distractors* and *type of information*.

Exhibit 13-17. Example of stimulus material for a relatively easy document literacy task

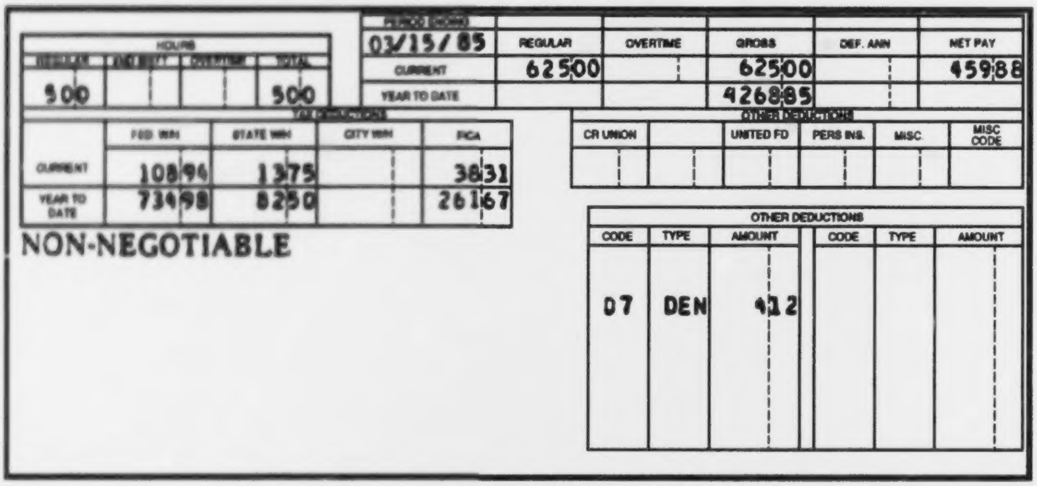
THEATER TRIP

A charter bus will leave from the bus stop (near the Conference Center) at 4 p.m., giving you plenty of time for dinner in New York. Return trip will start from West 45th street directly following the plays. Both theaters are on West 45th Street. Allow about 1 1/2 hours for the return trip.

Time:	**4 p.m., Saturday, November 20**		
Price:	**"On the Town"**	**Ticket and bus**	**$11**
	"Sleuth"	**Ticket and bus**	**$8.50**
Limit:	**Two tickets per person**		

A more difficult task at 275 on the document scale directs the reader to look at a wage and tax statement (Exhibit) and to select "gross pay for this year to date." If readers fail to identify and match on both features—gross pay and year to date—they are likely to respond with an incorrect amount based on distracting information. The structural complexity of this document was coded 5. It was rated 2 on *type of match* and *type of information* with a 3 for *plausibility of distractors*.

Exhibit 13-18. Example of stimulus material for moderately difficult document literacy tasks



HOURS

[illegible]	[illegible]	[illegible]	TOTAL
5 00			5 00

03/15/85	REGULAR	OVERTIME	GROSS	DEF. ANN	NET PAY
CURRENT	625 00		625 00		459 88
YEAR TO DATE			4268 85		

	FED WH	STATE WH	CITY WH	FICA
CURRENT	108 94	13 75		38 31
YEAR TO DATE	734 98	82 50		261 67

OTHER DEDUCTIONS

CR UNION		UNITED FD	PERS INS.	MISC.	MISC CODE

OTHER DEDUCTIONS

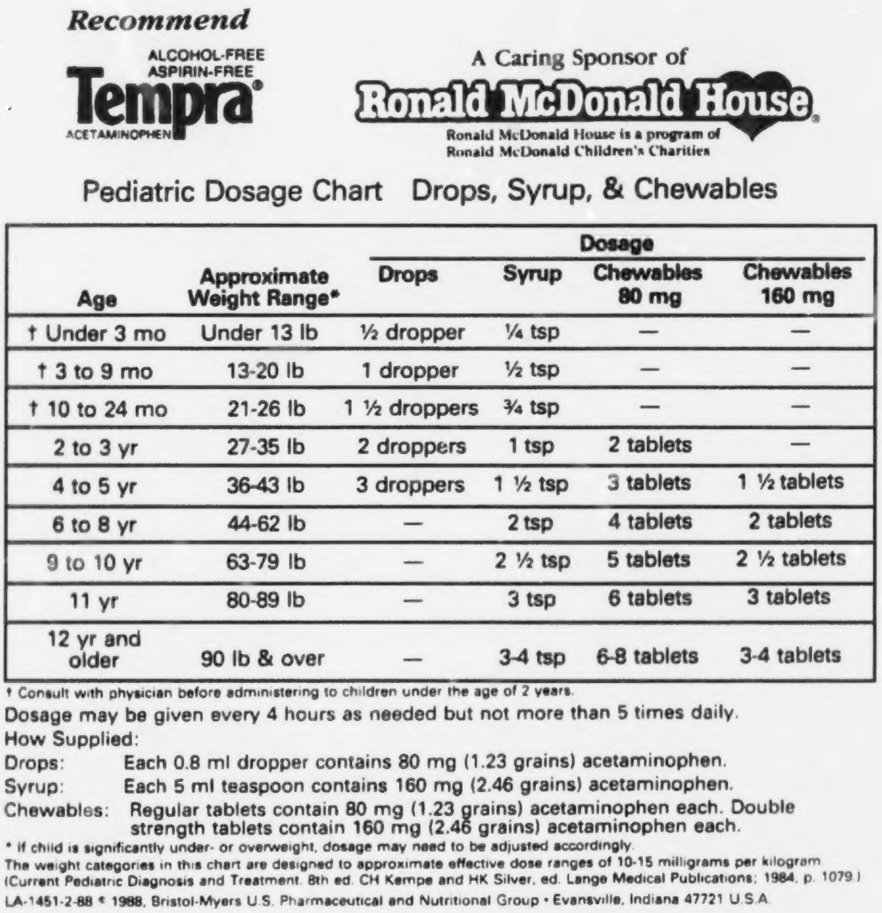
CODE	TYPE	AMOUNT	CODE	TYPE	AMOUNT
07	DEN	4 12			

NON-NEGOTIABLE

Another question using this wage and tax statement was also expected to require a two-feature match—current and net pay—and, therefore, to have approximately the same scale value. However, this task was considerably easier (224) and the codes assigned to the process variables indicate that little distracting information was present in the document. Each variable received a code value of 2.

Another task of similar difficulty (234) directs the reader to look at a pediatric dosage chart (Exhibit 13-19) and underline the sentence that indicates how often the medication may be administered. To respond successfully, the reader needs to associate the word "administered" in the directive to the word "given" in the document by looking at information outside the table itself. The structural complexity of this document was coded a 5. While *type of match* was coded 2, both *plausibility of distractors* and *type of information* received codes of 3.

Exhibit 13-19. Example of stimulus for moderately difficult document tasks



Recommend

ALCOHOL-FREE
ASPIRIN-FREE
Tempra®
ACETAMINOPHEN

A Caring Sponsor of
Ronald McDonald House®
Ronald McDonald House is a program of
Ronald McDonald Children's Charities

Pediatric Dosage Chart   Drops, Syrup, & Chewables

		Dosage			
Age	**Approximate Weight Range***	**Drops**	**Syrup**	**Chewables 80 mg**	**Chewables 160 mg**
† Under 3 mo	Under 13 lb	½ dropper	¼ tsp	—	—
† 3 to 9 mo	13-20 lb	1 dropper	½ tsp	—	—
† 10 to 24 mo	21-26 lb	1 ½ droppers	¾ tsp	—	—
2 to 3 yr	27-35 lb	2 droppers	1 tsp	2 tablets	—
4 to 5 yr	36-43 lb	3 droppers	1 ½ tsp	3 tablets	1 ½ tablets
6 to 8 yr	44-62 lb	—	2 tsp	4 tablets	2 tablets
9 to 10 yr	63-79 lb	—	2 ½ tsp	5 tablets	2 ½ tablets
11 yr	80-89 lb	—	3 tsp	6 tablets	3 tablets
12 yr and older	90 lb & over	—	3-4 tsp	6-8 tablets	3-4 tablets

† Consult with physician before administering to children under the age of 2 years.

Dosage may be given every 4 hours as needed but not more than 5 times daily.

How Supplied:

Drops: Each 0.8 ml dropper contains 80 mg (1.23 grains) acetaminophen.

Syrup: Each 5 ml teaspoon contains 160 mg (2.46 grains) acetaminophen.

Chewables: Regular tablets contain 80 mg (1.23 grains) acetaminophen each. Double strength tablets contain 160 mg (2.46 grains) acetaminophen each.

* If child is significantly under- or overweight, dosage may need to be adjusted accordingly.

The weight categories in this chart are designed to approximate effective dose ranges of 10-15 milligrams per kilogram (Current Pediatric Diagnosis and Treatment. 8th ed. CH Kempe and HK Silver, ed. Lange Medical Publications; 1984, p. 1079.)

LA-1451-2-88 © 1988, Bristol-Myers U.S. Pharmaceutical and Nutritional Group • Evansville, Indiana 47721 U.S.A.

(c) 1988, Bristol-Myers Pharmaceutical and Nutritional Group.

A second and more difficult question (327) using the medicine dosage chart directs the reader to determine from the chart how much syrup is recommended for a child who is 10 years old and weighs 50 pounds. This task is difficult because one can not simply match literal or synonymous information to perform successfully since the weight as given in the question is less than that of the typical 10 year old according to the table. Instead, one must rely on prior knowledge, or to find the asterisked note relating to the column headed "Approximate Weight Range," that the correct dosage is to be based on weight not age to ensure that the child receives an effective dose. In any event, if the reader approaches this task as a single literal match, the age of the child is a highly plausible distractor and may lead to an incorrect response. The variable codes reflect this line of reasoning with ratings of 4 for *type of match* and *plausibility of distractors* while *type of information* is rated 2.

13.4.2.3 Coding the 1985 and 1990 document literacy tasks

The preceding section provided examples to illustrate how type of match, type of information, and plausibility of distractors were coded for selected document literacy tasks from the 1990 literacy survey of job-seekers served by the Department of Labor. In this section, the coding rules are formulated in more general terms. A number of criteria must be taken into account when measuring the four variables associated with task difficulty on the document literacy scale.

13.4.2.3.1 Type of Match

Description. Type of match refers to the processes used to relate requested information to the corresponding information in a document, and to the process of entering a response. As with the prose literacy scale, locate, cycle, integrate, and generate tasks can be distinguished. Locate tasks on the document literacy scale require users to match one or more features in a question to one or more features in the document (Kirsch and Mosenthal, 1992a; Mosenthal, 1996). Based on this match, the answer is located in the appropriate node of a document. In a document, a node is either an element of a list, a matrix cell defined by the intersection of two or more lists, or a list itself. Cycle tasks, like those on the prose literacy scale, require users to perform an iterative series of locate matches and may involve the selection of several pieces of information that meet a criterion. With documents, cycle tasks are made difficult depending upon whether they are performed within a given list or between lists. Integrate tasks require users to compare or contrast information that has been located in two or more different locate matches or in one or more cycle matches (Kirsch and Mosenthal, 1992). Generate tasks require readers to use prior knowledge (often representing a specific type of content knowledge) to match information in a question or directive to corresponding information in a document (Mosenthal and Kirsch, 1993a, 1998).

Scoring rules. The scoring of type of match is basically determined by the nature of the match, and to a lesser extent by the number of phrases or features in the request, the number of responses requested, whether prior knowledge is needed for the required inferences, and how a choice among one of several possible answers should be selected.

Exhibit 13-20. Basic scoring rules for type of match: Nature of the task

Rule	Score
When the task is to *locate* the information in the document that corresponds to the features requested.	1
When the task is to *cycle* (that is, perform an iterative series of locate matches) to find the information that corresponds to the features requested. Add 1 point if the answer identified in one match is used to carry out a second match.	2
When the task is to *integrate* information located in a document by comparing	3
When the task is to integrate information located in a document by contrasting	4
When the task is to *generate* new information (that is, to use prior knowledge to match information requested with that in the document).	5

Sometimes matching is made more difficult as the number of *phrases or features* in the directions required to locate an answer increases (Kirsch and Mosenthal, 1990; Mosenthal, 1996). Matches that require the identification of only a single phrase or feature are, on average, easier than matches that require the identification of two phrases or features.

Exhibit 13-21. Additional scoring rule for type of match: Number of phrases or features in request

Rule	Add
When the request for information consists of two features	1
When the request for information consists of three features	2
When the request for information consists of four or more features	3

Sometimes matching is also made more difficult as the number of *responses* readers must supply increases and as the specificity of this number decreases. Requests of readers to list only one answer are easier than requests to list two or three answers; requests of readers to list two or three answers are easier than requests for four answers. Requests that do not specify the number of responses explicitly are harder than those that specify the number of multiple responses required.

Exhibit 13-22. Additional scoring rule for type of match: Number of responses requested

Rule	Add
When readers are requested to list two or three answers	1
When readers are requested to list four or five answers $ add 1 point if the request does not specify exactly how many	2
When readers are requested to list six or more answers $ add 1 point if the request does not specify exactly how many	3

Matching can be made even more difficult to the extent that readers have to *make inferences* to match information in the question to information in the document (Mosenthal and Kirsch, 1993b, 1998). Questions may require either low text-based inferences, which consist of inferences which can be made within the context of information provided in a text, or high text-based inferences, which consist of inferences which require some combination of knowledge of the text and specialized prior knowledge.

Exhibit 13-23. Additional scoring rule for type of match: Inferences needed

Rule	Add
When the text alone provides sufficient information to make an inference needed to match the request with the information in the text or document (a low text-based inference)	1
When prior knowledge as well as the text is needed to make the inference needed to match the request with the information in the text or document (a high text-based inference)	3

Sometimes matching is made more difficult when all possibilities match the request and readers have to choose which one of several possible answers best completes a requested *information frame* (Mosenthal and Kirsch, 1991, 1996). In these cases, the match is more difficult when this choice requires using the text to infer why one of several possible answers best completes a requested information frame, when this choice requires identifying conditional information which renders one of the possible answers more consistent with the conditions requested than others, or when this choice requires readers to relate a pronoun to its antecedent before an answer can be provided.

Exhibit 13-24. Additional scoring rule for type of match: Completing an information frame

Rule	Add
When the choice among candidate answers requires a low, text-based inference (the text alone provides sufficient information to make the inference), the identification of a condition, the identification of a pronoun antecedent, or a restatement of a type of information	2
When the choice among candidate answers requires a high, text-based inference (prior knowledge as well as the text is needed to make the inference)	4

These scoring rules are additive (Meyer, Marsiske, and Willis, 1993). A document literacy task, for another example, might have a basic score of 1 because it is a locate task, but have additional points added because two simultaneous features must be matched (add 1), and the answer should consist of three responses (add another 1) whose actual number is also not explicitly specified (add 1). A document assessment task with these features would have a total type-of-match score of 4.

While the scoring system for type of match could theoretically generate scores as high as 20, this was not the case with the 1985 and 1990 document literacy tasks. The actual document-based tasks used scores from 1 to 8 on type of match. Tasks beyond this level, while possible, would be so difficult that they would bear little resemblance to the ordinary usage of documents.

13.4.2.3.2 Plausibility of Distractors

Description. Distractors are elements of a question's given or requested information that appear in the prose text or document, but when identified lack an essential feature and do not qualify as correct. Unless the possible but incorrect answers share some of the features of the correct answer, they do not function as distractors. Defining how close the distractor is to the correct answer involves a series of decision rules that differ somewhat depending on whether the material occurs in the form of prose texts or documents (Mosenthal, 1996, 1998).

Scoring rules. Tasks are easiest when the document contains no information related to the conditions set forth in the question other than the answer. Tasks are also easiest when there is only a single item in a list, or there is only one list with a unique label unrelated to the other labels in a document.

Tasks become slightly more difficult when a distractor appears, but does not occur very close to the correct answer. This occurs in documents when there is more than one item in a list in which one is searching for requested information or when there are labels in other lists that bear a resemblance to the label on which one is searching.

Tasks become more difficult when plausible distractors for both given and requested information appear, but are not both located next to the correct information. This occurs in documents when they appear in different matrix cells or in lists other than the cell or list in which an answer actually appears.

Tasks become still more difficult when plausible distractors for both given and requested information appear in the same place but are not located near the correct information. This occurs in documents when one or more features from both appear in a matrix cell or list other than the one in which the answer appears.

Tasks are most difficult when plausible distractors for both given and requested information appear in the same place, or appear with negative conditions, and are located near the correct information. This occurs in documents when one or more features from both requested and given information appear in the same matrix cell or list as the answer.

Exhibit 13-25. Scoring rules for plausibility of distractors

Rule	Score
When no information related to the conditions requested appears, other than the answer (no plausible distractors)	1
When other information somewhere in a document (but not near the answer) bears a resemblance to the requested information	2
When distractors for both given and requested information appear in different matrix cells or in lists other than the cell or list containing an answer	3
When distractors for both given and requested information appear in a matrix cell or list other than the answer node	4
When distractors for both given and requested information appear in the same matrix cell or list as the answer	5

13.4.2.3.3 Type of Information

Description. Type of information refers to the degree of concreteness of the objects described in a prose text or document (Mosenthal and Kirsch, 1991b). More abstract objects are harder for readers to identify and understand.

Scoring rules. The scoring rules for type of information were identical for prose texts and documents, so the same procedures were used.

Exhibit 13-26. Scoring rules for type of information

Rule	Score
When the information requested refers to a person, animal, place, or thing (most concrete)	1
When the information requested refers to an amount, time, attribute, type, action, or location	2
When the information requested refers to a manner, goal, purpose, alternative, attempt, condition, pronominal reference, or predicate adjective	3
When the information requested refers to a cause, effect, reason, result, evidence, similarity, or explanation	4
When the information requested refers to an equivalence, difference, theme, or pattern (most abstract)	5

13.4.2.3.4 Structural Complexity

Description. Just as prose texts vary in the length of sentences, number of syllables in the words used, and the complexity of the syntax, documents also vary in their complexity. The complexity of documents depends on their organization, number of elements, and number of labels. The measurement of the complexity of documents derives from research by Mosenthal and Kirsch (1989, 1991a, 1998) and their predecessors (Meyer and Rice, 1984).

Scoring rules for documents. Structural complexity of documents is based on the type of document, along with the number of items and labels comprising the document. The basic structural unit of documents is the simple list, which consists of a series of items or elements that belong to a common, more generic category called a "label." More complex documents are combinations of simple lists: combined, intersected, nested, or multiple lists, depending on their relationships to one another. After the initial score is derived from the document type, a supplementary set of rules is applied concerning the number of items and labels comprising the document.

Exhibit 13-27. Basic scoring rules for structural complexity: Type of document

Rule	Score
When the document has a simple list structure.	1
When the document has a combined list structure.	2
When the document has an intersected list structure.	3
When the document has a nested list structure.	4
When the document consists of different multiple documents or a combined list with minimal structure.	5

Next, the items in each list are counted. In most cases, items are the cells or basic elements of any given list, but for lists that consist of a series of sentences, each independent and dependent clause within each sentence is counted as a separate item. As shown in the table below, the additional credit to be added to the basic score depends on the number of items.

Exhibit 13-28. Additional scoring rule for structural complexity: Number of items in lists

Rule	Add
When the lists consist of 76-125 items	1
When the lists consist of 126-175 items	2
When the lists consist of more than 175 items	3

Finally, the labels heading each list are counted. Labels are the list headings that describe the contents of the list. Complex documents containing more than one list can have many labels for different parts of the document. As shown in the table below, the additional credit to be added to the basic score depends on the number of labels.

Exhibit 13-29. Additional scoring rule for structural complexity: Number of labels in lists

Rule	Add
When the document contains 16-25 labels	1
When the document contains 26-35 labels	2
When the document contains more than 35 labels	3

These scoring rules are additive. For example, a document might be assigned a structural complexity score of 2 because it is a combined list (begin with 2 points) which involves 100 items (add 1) in 3 labeled columns (don't add anything). The total structural complexity score, given the combined features of the document, is 3.

13.4.2.4 Codes for all 1985 and 1990 document literacy tasks

The preceding sections described in detail the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the document literacy scale. These rules were applied to all document literacy tasks in the 1990 survey and in the 1985 young adult literacy assessment, and the resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-4.

Table 13-4. List of document literacy tasks, along with RP80 task difficulty score, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractor, type of information): 1990 survey of the literacy of job-seekers

Identifier	Task Description	RP80	IRT parameters			Complexity	Match	Distractor	Information
			a	b	c				
SCOR100	Social Security card: Sign name on line	70	0.505	-4.804	0.000	1	1	1	1
SCOR300	Driver's license: Locate expiration date	152	0.918	-2.525	0.000	2	1	2	1
SCOR200	Traffic Signs	176	0.566	-2.567	0.000	1	1	1	1
AB60803	Nurses' convention: What is time of program?	181	1.439	-1.650	0.000	1	1	1	1
AB60802	Nurses' convention: What is date of program?	187	1.232	-1.620	0.000	1	1	1	1
SCOR400	Medicine dosage	186	0.442	-2.779	0.000	2	1	2	2
AB71201	Mark correct movie from given information	189	0.940	-1.802	0.000	8	2	2	1
A110501	Registration & tuition Info	189	0.763	-1.960	0.000	3	1	2	2
AB70104	Job application: Complete personal information	193	0.543	-2.337	0.000	1	2	1	2
AB60801	Nurses' convention: Write correct day of program	199	1.017	-1.539	0.000	1	1	2	1
SCOR500	Theatre trip information	197	0.671	-1.952	0.000	2	1	2	2
AB60301	Phone message: Write correct name of caller	200	1.454	-1.283	0.000	1	1	2	1
AB60302	Phone message: Write correct number of caller	202	1.069	-1.434	0.000	1	1	1	1
AB80301	How companies share market	203	1.292	-1.250	0.000	7	2	2	2
AB60401	Food coupons	204	0.633	-1.898	0.000	3	2	2	1
AB60701	Nurses' convention: Who would be asked questions	206	1.179	-1.296	0.000	1	2	2	1
A120601	MasterCard/Visa statement	211	0.997	-1.296	0.000	6	1	2	2
AB61001	Nurses' convention: Write correct place for tables	217	0.766	-1.454	0.000	1	1	2	2
A110301	Dessert recipes	216	1.029	-1.173	0.000	5	3	2	1
AB70903	Checking deposit: Enter correct amount of check	223	1.266	-0.922	0.000	3	2	2	1
AB70901	Checking deposit: Enter correct date	224	0.990	-1.089	0.000	3	1	1	1
AB50801	Wage & tax statement: What is current net pay?	224	0.734	-1.366	0.000	5	2	2	2
A130201	El Paso Gas & Electric bill	223	1.317	-0.868	0.000	8	1	2	2
AB70801	Classified: Match list with coupons	229	1.143	-0.881	0.000	8	2	3	1
AB30101	Street map: Locate intersection	232	0.954	-0.956	0.000	4	2	2	2
AB30201	Sign out sheet: Respond to call about resident	232	0.615	-1.408	0.000	2	3	2	1
AB40101	School registration: Mark correct age information	234	0.821	-1.063	0.000	6	2	2	3
A131201	Tempra dosage chart	233	1.005	-0.872	0.000	5	2	3	3
AB31301	Facts about fire: Mark information in article	235	0.721	-1.170	0.000	1	2	3	2
AB80401	How companies share market	236	1.014	-0.815	0.000	7	3	2	2
AB60306	Phone message: Write whom message is for	237	0.948	-0.868	0.000	1	2	3	1
AB60104	Make out check: Enter correct amount written out	238	1.538	-0.525	0.000	6	3	2	1
AB21301	Bus schedule	238	0.593	-1.345	0.000	2	2	3	2
A110201	Dessert recipes	239	0.821	-0.947	0.000	5	3	2	1
AB30301	Sign out sheet: Respond to call about resident	240	0.904	-0.845	0.000	2	2	2	3
AB[illegible]0701	Major medical:locate Eligibility from table	245	0.961	-0.703	0.000	4	2	2	2
AB60103	Make out check: Enter correct amount in numbers	245	0.993	-0.674	0.000	6	3	2	1
AB60101	Make out check: Enter correct date on check	246	1.254	-0.497	0.000	6	3	2	1
AB60102	Make out check: Paid to the correct place	246	1.408	-0.425	0.000	6	3	2	1
AB50401	Catalog order: Order product one	247	0.773	-0.883	0.000	8	3	2	1
AB60303	Phone message: Mark "please call" box	249	0.904	-0.680	0.000	1	2	2	2
AB50701	Almanac football: Explain why an award is given	254	1.182	-0.373	0.000	6	2	2	3
AB20101	Energy graph: Find answer for given conditions (1)	255	1.154	-0.193	0.228	4	3	2	1
A120901	MasterCard/Visa statement	257	0.610	-0.974	0.000	6	1	2	2
A130101	El Paso Gas & Electric bill	257	0.953	-0.483	0.000	8	2	2	2
AB91101	Minimum wage power	260	0.921	-0.447	0.000	4	3	3	2
AB81001	Consumer Reports books	261	1.093	-0.304	0.000	4	3	2	1
AB90101	Pest control warning	261	0.889	-0.471	0.000	2	3	3	2
AB21501	With graph, predict sales for spring 1985	261	0.799	-0.572	0.000	5	3	2	2
AB20601	Yellow pages: Find place open Saturday	266	1.078	-0.143	0.106	7	3	2	1
A130401	El Paso Gas & Electric bill	270	0.635	-0.663	0.000	8	3	3	2
AB70902	Checking deposit: Enter correct cash amount	271	0.858	-0.303	0.000	3	3	3	2

Table 13-4. List of document literacy tasks, along with RP80 task difficulty score, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractor, type of information): 1990 survey of the literacy of job-seekers — Continued

Identifier	Task Description	RP80	IRT parameters a	IRT parameters b	IRT parameters c	Complexity	Match	Distractor	Informatio
AB50601	Almanac football: Locate page of info in almanac	276	1.001	-0.083	0.000	5	3	2	2
A110701	Registration & tuition Info	277	0.820	-0.246	0.000	3	2	5	2
AB20201	Energy graph: Find answer for given conditions (2)	278	0.936	-0.023	0.097	4[3]	4[3]	2[3]	1[3]
AB31101	Abrasive gd: Can product be used in given case?	280	0.762	-0.257	0.000	10	5	2	3
AB80101	Burning out of control	281	0.550	-0.656	0.000	2	3	2	2
AB70701	Follow directions on map: Give correct location	284	0.799	-0.126	0.000	4	4	2	2
A110801	Washington/Boston schedule	284	0.491	-0.766	0.000	9	2	4	2
AB70301	Almanac vitamins: Locate list of info in almanac	287	0.754	-0.134	0.000	5	3	4	2
AB20401	Yellow pages: Find a list of stores	289	0.479	-0.468	0.144	7	2	5	1
AB20501	Yellow pages: Find phone number of given place	291	0.415	-0.772	0.088	7	2	4	2
AB60305	Phone message: Write who took the message	293	0.640	-0.221	0.000	1	5	2	1
AB30401	Sign out sheet: Respond to call about resident (2)	297	0.666	-0.089	0.000	2	2	1	4
AB31001	Abrasive guide: Type of sandpaper for sealing	304	0.831	0.285	0.000	10	4	2	2
AB20301	Energy: Yr 2000 source prcnt power larger than 71	307	1.090	0.684	0.142	4	4	2	1
AB90901	U.S. Savings Bonds	308	0.932	0.479	0.000	6	4	4	2
AB60304	Phone message: Write out correct message	310	0.895	0.462	0.000	1	5	2	3
AB81002	Consumer Reports books	311	0.975	0.570	0.000	4	3	5	2
AB20801	Bus schd: Take correct bus for given condition (2)	313	1.282	0.902	0.144	10	3	5	2
AB50402	Catalog order: Order product two	314	1.108	0.717	0.000	8	4	4	3
AB40401	Almanac: Find page containing chart for given info	314	0.771	0.397	0.000	5	4	3	2
AB21001	Bus schd: Take correct bus for given condition (4)	315	0.730	0.521	0.144	10	3	4	2
AB60502	Petroleum graph: Complete graph including axes	318	1.082	0.783	0.000	10	6	2	2
A120701	MasterCard/Visa statement	320	0.513	-0.015	0.000	6	2	4	2
AB20701	Bus schd: Take correct bus for given condition (1)	324	0.522	0.293	0.131	10	3	4	2
A131301	Tempra dosage chart	326	0.624	0.386	0.000	5	4	4	2
AB50501	Telephone bill: Mark information on bill	330	0.360	-0.512	0.000	7	4	4	2
AB91401	Consumer Reports index	330	0.852	0.801	0.000	7	3	5	3
AB30801	Almanac: Find page containing chart for given info	347	0.704	0.929	0.000	5	4	5	2
AB20901	Bus schd: After 2:35, how long til Flint&Acad bus	348	1.169	1.521	0.163	10	5	4	2
A130301	El Paso Gas & Electric bill	362	0.980	1.539	0.000	8	5	4	5
A120801	MasterCard/Visa statement	363	0.727	1.266	0.000	6	5	4	2
AB91301	Consumer Reports index	367	0.620	1.158	0.000	7	4	5	3
AB60501	Petroleum graph: Label axes of graph	378	1.103	1.938	0.000	11	7	2	5
AB30901	Almanac: Determine pattern in exports across years	380	0.299	0.000	0.000	7	5	5	3
A100701	Spotlight economy	381	0.746	1.636	0.000	10	5	5	2
A100501	Spotlight economy	386	0.982	1.993	0.000	10	5	5	5
A100401	Spotlight economy	406	0.489	1.545	0.000	10	5	5	2
AB51001	Income tax table	421	0.257	0.328	0.000	9	4	5	2
A100601	Spotlight economy	465	0.510	2.737	0.000	10	7	5	2

13.4.2.5 Validity evidence for the document scale

As with the prose tasks, an important piece of validation evidence concerns the document tasks newly developed for the 1990 literacy survey of job-seekers served by the Department of Labor. These new tasks were designed to reflect various aspects of the theoretical framework as it evolved from the 1985 young adult literacy assessment. Table 13-5 gives the results of regression analyses for the 1990 document literacy tasks as well as for the combined set (including the 1985 document tasks). Overall, the variance accounted for reached 92 percent for the new tasks and 87 percent for the combined set of the 1985 and 1990 document literacy tasks.

Table 13-5. Unstandardized regression coefficients and standard errors predicting RP80 task difficulties on the basis of four structure and process variables: 33 new document literacy tasks and 92 total document literacy tasks from the 1990 survey of job-seekers

	New Tasks			All Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural complexity	5.17	1.91	.01	1.39	1.10	.21
Process Variables						
Type of match	24.12	3.70	.00	24.46	2.28	.00
Plausibility of distractors	23.84	4.11	.00	22.71	2.44	.00
Type of information	-1.35	4.93	.79	9.09	3.15	.00
Variance accounted for:						
R^2		92%			87%	
Adjusted R^2		91%			86%	
Degrees of freedom		28			87	

13.4.3 Quantitative Literacy

Since adults are often required to perform numerical operations in everyday life, the ability to perform quantitative tasks is an important area of adult literacy. To complete these types of tasks successfully, a respondent must perform arithmetic operations such as addition, subtraction, multiplication, or division either singly or in combination using numbers or quantities that are embedded in printed information.

At first glance, quantitative tasks might appear to represent fundamentally different skills from those involved in processing prose and documents. However, an analysis of tasks along this scale shows that the difficulty of these quantitative tasks is affected by the processing of the printed information in which they are contained and thus by the processing variables salient for prose and document tasks.

13.4.3.1 Quantitative variables

In general, it appears that many individuals can perform simple arithmetic operations when both the numbers and operations are made explicit. Yet, when these same operations are performed on numbers that must be located and extracted from different types of documents that contain similar but irrelevant information, or when these operations must be inferred from printed directions, quantitative tasks become increasingly difficult. To complete tasks on the quantitative scale, individuals are required to match information in a question or directive to information stated in one or more documents or pieces of text. In addition, tasks from the quantitative scale may require the reader to negotiate information that can serve as *plausible distractors* during the calculation of a correct response. Moreover, individuals are also required to process some type of printed information. While *type of information* varied for prose and documents, requested information for the quantitative tasks is always an amount. The stimulus materials for the quantitative tasks are mostly documents, and these vary widely in their structural complexity.

Thus, while the quantitative tasks include *structural complexity*, *type of match* and *plausibility of distractors* as defined for the prose and document tasks, they also involve two "formulate" variables that are unique to this scale. The first formulate variable, *operation specificity*, refers to the process of identifying (and sometimes entering) the numbers in an arithmetic expression, including the determination of the appropriate operation(s) that must be performed. Tasks tend to be more difficult when the numbers must be identified in a document and when these numbers are not in column format or adjacent to each other. Tasks also tend to become more difficult when the operation is not specified or when the wording in the question or directive does not contain an explicit semantic relation statement such as "how many" or "calculate the difference." This variable was coded from 1 (easiest) to 9 (most difficult) based on a set of additive rules reflecting the various facets stated here.

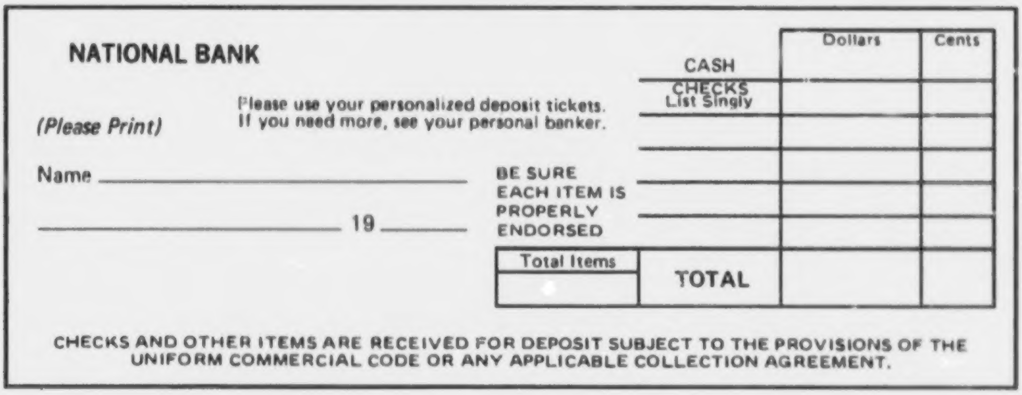
The second formulate variable, *type of calculation*, includes both the type of arithmetic operation (addition, subtraction, multiplication, or division) required to produce a correct response, and whether that operation must be performed alone or in combination. Tasks requiring two or more operations tend to be more difficult than those involving a single operation. This variable ranged from 1 (easiest) to 5 (most difficult).

13.4.3.2 Examples of quantitative literacy tasks

The least demanding quantitative task in the 1990 survey (220) required the reader to enter and total two numbers on a bank deposit slip (Exhibit 13-30). In this example, both the number and operation were judged to be easily identified, and the operation involved the simple addition of two decimal numbers that were presented in column format. Moreover, the numbers were stated in the directive so that the problem was, in some sense, set up for the reader. As a result, each of the process and formulate variables received a code of 1. The structural complexity of the document was coded 2.

In other tasks having similar characteristics that received somewhat higher values on the scale, the quantities, while easy to identify, were not explicitly given in the directive but had to be searched for and identified in the document. One such task having a scale value of 270 required the reader to locate the appropriate shipping charges in a table before entering the correct amount on an order form and calculating the total price for ordering office supplies. The structural complexity of this document was judged to be 6, while *type of match* was coded 3, and *plausibility of distractors* was coded 2. In addition, *type of calculation* received a code of 1 and *operation specificity* a code of 3.

Exhibit 13-30. Example of stimulus material for a relatively easy quantitative literacy task



NATIONAL BANK

(Please Print)

Please use your personalized deposit tickets.
If you need more, see your personal banker.

Name ______________________

______________________ 19 ______

BE SURE
EACH ITEM IS
PROPERLY
ENDORSED

	Dollars	Cents
CASH		
CHECKS List Singly		
Total Items ______ TOTAL		

CHECKS AND OTHER ITEMS ARE RECEIVED FOR DEPOSIT SUBJECT TO THE PROVISIONS OF THE UNIFORM COMMERCIAL CODE OR ANY APPLICABLE COLLECTION AGREEMENT.

Tasks around 300 on the quantitative scale still require a single arithmetic operation. What appears to distinguish these tasks, however, is the fact that the reader must identify, in various places in the document, two or more numbers needed to solve the problem. The numbers are not presented in column format, nor is the operation needed to complete the task explicitly stated in the directive or provided by the format of the document, as in the previous examples. Instead, the operation must be determined from arithmetic relation terms, such as "how many" or "what is the difference" used in the question.

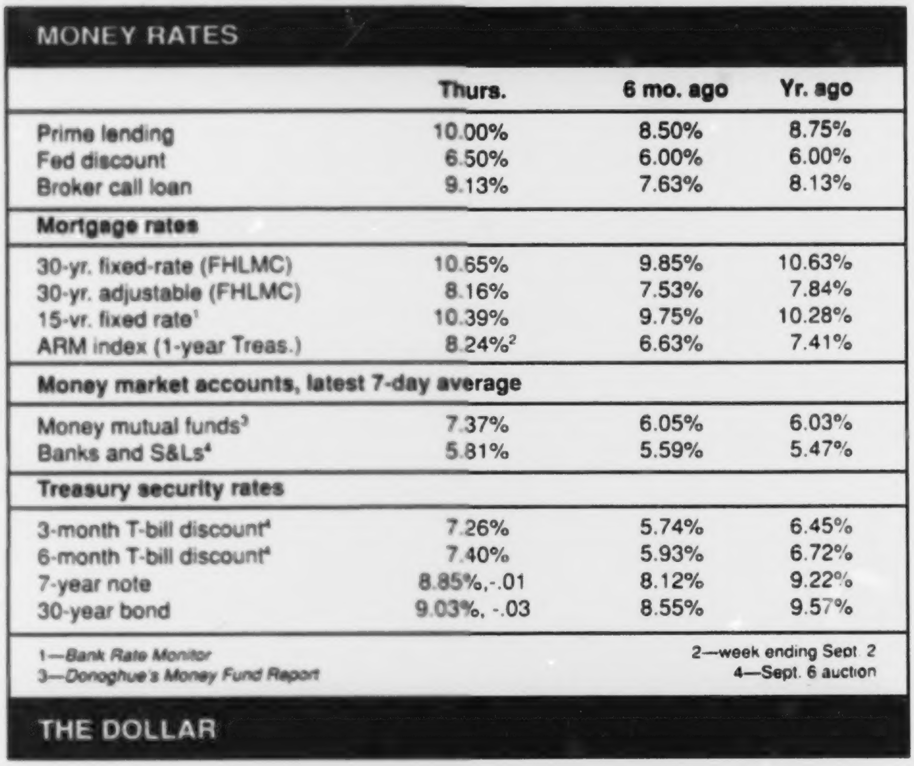
One such task receiving a scale value of 312 requires the reader to look at a table of money rates (Exhibit 13-31) to determine how much more interest would be earned in money market accounts provided by mutual funds than in those provided by S&Ls. This document received a structural complexity score of 4. It was also coded 3 for *type of match* and *plausibility of distractors* and 2 for each of the two formulate variables—*type of calculation* and *operation specificity*.

Tasks with the highest scale values (above 370) tended to require the reader to draw heavily on background information in order to identify both the quantities and the operations needed to complete the task successfully. For example, the most difficult quantitative task used in this assessment required readers to look at a newspaper advertisement for a home equity loan (Exhibit 13-32) and then, using the information provided, explain how they would calculate the total amount of interest charges to be paid. This document received a structural complexity score of 2. It was coded 5 for *type of match, plausibility of distractors*, and *type of calculation*, while *operation specificity* received a coded value of 7.

13.4.3.3 Coding the 1985 and 1990 Quantitative Literacy Tasks

The preceding section provided examples to illustrate how specificity of operation, type of calculation, and plausibility of distractors were coded for selected quantitative literacy tasks from the 1990 literacy survey of job-seekers served by the Department of Labor. In this section, the coding rules are formulated

Exhibit 13-31. Example of stimulus material for relatively difficult quantitative literacy task



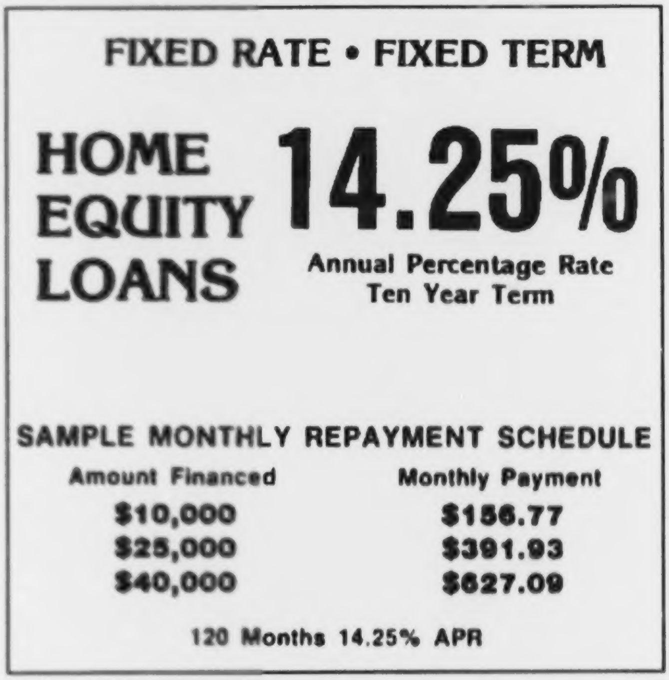
MONEY RATES

	Thurs.	6 mo. ago	Yr. ago
Prime lending	10.00%	8.50%	8.75%
Fed discount	6.50%	6.00%	6.00%
Broker call loan	9.13%	7.63%	8.13%
Mortgage rates			
30-yr. fixed-rate (FHLMC)	10.65%	9.85%	10.63%
30-yr. adjustable (FHLMC)	8.16%	7.53%	7.84%
15-yr. fixed rate[1]	10.39%	9.75%	10.28%
ARM index (1-year Treas.)	8.24%[2]	6.63%	7.41%
Money market accounts, latest 7-day average			
Money mutual funds[3]	7.37%	6.05%	6.03%
Banks and S&Ls[4]	5.81%	5.59%	5.47%
Treasury security rates			
3-month T-bill discount[4]	7.26%	5.74%	6.45%
6-month T-bill discount[4]	7.40%	5.93%	6.72%
7-year note	8.85%, -.01	8.12%	9.22%
30-year bond	9.03%, -.03	8.55%	9.57%

1—*Bank Rate Monitor*
3—*Donoghue's Money Fund Report*
2—week ending Sept. 2
4—Sept. 6 auction

THE DOLLAR

Exhibit 13-32. Example of stimulus material for a difficult quantitative literacy task



FIXED RATE • FIXED TERM

HOME EQUITY LOANS

14.25%

Annual Percentage Rate
Ten Year Term

SAMPLE MONTHLY REPAYMENT SCHEDULE

Amount Financed	Monthly Payment
$10,000	$156.77
$25,000	$391.93
$40,000	$627.09

120 Months 14.25% APR

in more general terms. A number of criteria must be taken into account when measuring the four variables associated with task difficulty on the quantitative literacy scale.

13.4.3.3.1 Specificity of Operation

Description. To obtain the requested quantitative information, readers must identify the relevant quantities, understand their relationships to one another, set up an arithmetic equation based on these relationships, and carry out simple arithmetic manipulations. Instructions for how to formulate an arithmetic expression can be more or less specific about identifying the relevant quantities and the relationships among them. Further, the operations necessary to obtain the requested information can be more or less specific.

Scoring rules. The evaluation of operation specificity takes into account aspects of both the necessary operation and the amounts involved. The *relationship* among the relevant amounts can be specified in terms that directly translate into an operation, that translate more indirectly, or that require a larger quantitative vocabulary to understand the relationship. Quantitative tasks are easier if the relationships among the amounts are described with arithmetic symbols or common arithmetic terms. Tasks are more difficult if the terminology describing the relationship uses a more specialized vocabulary or requires understanding ratios.

Exhibit 13-33. Basic scoring rule for specificity of required operation: Relationships of amounts

Rule	Score
When the operation is specifically identified by words or symbols, using terms such as 'add', 'total' or '+'; 'subtract' or '-'; 'multiply' or 'x'; and 'divide' or '÷'	0
When the operation is specified with a relationship among the numbers, using terms such as 'how much more'; 'how much less' or 'calculate the difference'; and 'how many times'	1
When the operation is identified using more specialized vocabulary, using terms such as 'how much is saved'; 'how much is the deduction'; or 'what is the net profit'	2
When the operation is identified as a unit ratio, such as 'miles per gallon', 'cost per square foot', or 'price per square yard'	3

Operation specificity is made more difficult as the *amounts* involved are harder to identify or harder to use in arithmetic operations. Quantitative tasks are easier if the amounts appear in a row or column format, if they are adjacent to one another, if they are labeled, if they do not require a search, if they are one-step problems, or if they do not involve conversions of units of measurement. Tasks are harder if they are not in a row or column format, if the amounts are not adjacent, if the labels associated with the amounts have to be inferred, if the amounts require a search, if they involve more than one step, or if they require converting units of measurement.

Exhibit 13-34. Additional scoring rule for specificity of required operation: Identifying amounts

Rule	Add
If the amounts are • in a row and column format • adjacent to one another • presented in the current task and no search is needed, if the problem requires a single step, if the labels are present and no inference is needed, or if the amounts need not be transformed into common units	0
If the amounts are *not* in a row and column format	1
If the amounts are *not* adjacent to one another	1
If the amounts are *not* presented, but must be carried over from a prior task, or the problem requires more than one step	1
If the amounts are *not* presented, but must be identified by a search	1
If the labels for the amounts must be inferred	1
If the amounts are in different units (such as time in hours and minutes or fractions with different denominators) that must be transformed into a common unit	1

The scoring rules for operation specificity are additive and scored by adding a point to the basic score for any of several possible factors that can make the relevant amounts more difficult to identify and manipulate. A quantitative task, for example, might have a basic score of 1 because it specifies with common terminology a relationship among the numbers, but has additional points added because the amounts are not in a row and column format (add 1) nor are they adjacent to one another (add 1), and the amounts must be transformed into a common unit of measurement (add 1). A quantitative assessment task with these features would have a total operation-specificity score of 4.

13.4.3.3.2 Type of Calculation

Description. Type of calculation measures the complexity of the various operations that readers use to relate one set of numbers to another in order to produce a sum, difference, product, or quotient. Quantitative tasks are easiest when the calculation is a single sum and most difficult when more than one of these calculations is required.

Scoring rules. For tasks that involve a single operation, those that involve addition are the easiest; those that involve subtraction are next easiest; those that involve multiplication are more difficult; and those involve division are the most difficult. When the reader must manipulate numbers that are the outcome of operations in preceding tasks, then the job becomes even more difficult. Any task that requires two or more operations (such as a division followed by a multiplication) is more difficult than any that requires only a single operation.

Exhibit 13-35. Scoring rules for type of calculation

Rule	Score
When two quantities are to be added (easiest)	1
When two quantities are to be subtracted	2
When two quantities are to be multiplied	3
When one quantity is to be divided by another	4
When readers are requested to operate on two quantities and use the result with another quantity to perform a second operation (hardest)	5

13.4.3.3.3 Type of Match, Plausibility of Distractors, and Structural Complexity

The task features type of match and plausibility of distractors apply to quantitative literacy tasks in the same way as they did to prose and document literacy tasks. Structural complexity is a feature of the prose text or documents in which the quantitative information is embedded. It was also defined in the same way as it was for the prose and document literacy tasks.

13.4.3.4 Codes for all 1985 and 1990 quantitative literacy tasks

The preceding sections described in detail the several criteria that must be taken into account when measuring the variables associated with task difficulty on the quantitative literacy scale. These rules were applied to all quantitative literacy tasks in the 1990 survey and in the 1985 young adult literacy assessment. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-6.

Table 13-6. List of quantitative literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractors, type of calculation, and specificity of operation): 1990 survey of the literacy of job-seekers

Identifier	Quantitative Literacy Items	RP80	IRT parameters			Complexity	Match	Distractor	Calculation	Op specfy
			a	b	c					
AB70904	Enter total amount of both checks being deposited	221	0.869	-1.970	0.000	2	1	1	1	1
AB50404	Catalog order: Shipping, handling, and total	271	0.968	-0.952	0.000	6	3	2	1	3
AB91201	Tempra coupon	271	0.947	-0.977	0.000	1	2	1	5	4
AB40701	Check ledger: Complete ledger (1)	277	1.597	-0.501	0.000	3	2	2	1	4
A121001	Insurance protection workform	275	0.936	-0.898	0.000	2	3	2	3	2
AB90102	Pest control warning	279	0.883	-0.881	0.000	2	3	3	1	4
AB40702	Check ledger: Complete ledger (2)	281	1.936	-0.345	0.000	3	2	2	2	4
AB40703	Check ledger: Complete ledger (3)	282	1.874	-0.332	0.000	3	1	2	2	4
A131601	Money rates: Thursday vs. one year ago	281	1.073	-0.679	0.000	4	3	2	2	4
AB40704	Check ledger: Complete ledger (4)	283	1.970	-0.295	0.000	3	2	2	2	4
AB80201	Burning out of control	286	0.848	-0.790	0.000	2	3	2	2	4
A110101	Dessert recipes	289	0.813	-0.775	0.000	5	3	2	2	4
AB90201	LPGA money leaders	294	0.896	-0.588	0.000	5	2	2	2	4
A120101	Businessland printer stand	300	1.022	-0.369	0.000	2	3	3	2	4
AB81003	Consumer Reports books	301	0.769	-0.609	0.000	7	2	3	1	4
AB80601	Valet airport parking discount	307	0.567	-0.886	0.000	2	3	3	2	4
AB40301	Unit price: Mark economical brand	311	0.816	0.217	0.448	2	2	3	4	6
A131701	Money rates: compare S&L w/ Mutual funds	312	1.001	-0.169	0.000	4	3	3	2	2
AB80701	Valet airport parking discount	315	0.705	-0.450	0.000	2	2	3	3	4

Table 13-6. List of quantitative literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractors, type of calculation, and specificity of operation): 1990 survey of the literacy of job-seekers — Continued

Identifier	Quantitative Literacy Items	RP80	IRT parameters			Complexity	Match	Distractor	Calculation	Op specfy
			a	b	c					
A100101	Pizza coupons	316	0.690	-0.472	0.008	2	3	3	1	4
AB90301	LPGA money leaders	320	1.044	0.017	0.000	5	1	2	4	3
A110401	Dessert recipes	323	1.180	0.157	0.000	5	3	2	3	6
A131401	Tempra dosage chart	322	1.038	0.046	0.000	5	3	3	2	4
AB40501	Airline schedule: plan travel arrangements (1)	326	0.910	0.006	0.000	3	3	3	5	3
AB70501	Lunch: Determine correct change using info in menu	331	0.894	0.091	0.000	2	2	2	5	4
A120201	Businessland printer stand	340	0.871	0.232	0.000	2	3	4	3	5
A110901	Washington/Boston train schedule	340	1.038	0.371	0.000	7	4	4	2	5
AB60901	Nurses Convention: Write number of seats needed	346	0.504	-0.355	0.000	3	4	4	1	5
AB70601	Lunch: Determine 10% tip using given info	349	0.873	0.384	0.000	2	1	2	5	7
A111001	Washington/Boston train schedule	355	0.815	0.434	0.000	7	4	4	2	5
A130501	El Paso Gas & Electric bill	352	0.772	0.323	0.000	8	3	4	2	2
A100801	Spotlight economy	356	0.874	0.520	0.000	8	5	4	2	2
AB40201	Unit price: Estimate cost/oz of peanut butter	356	0.818	0.455	0.000	2	1	2	4	5
A121101	Insurance protection workform	356	0.860	0.513	0.000	2	1	2	5	4
A100901	Camp advertisement	366	0.683	0.447	0.000	2	2	4	5	4
A101001	Camp advertisement	366	0.974	0.795	0.000	2	3	4	5	4
AB80501	How companies share market	371	1.163	1.027	0.000	6	3	2	3	6
A131501	Tempra dosage chart	381	0.916	1.031	0.000	5	3	5	3	5
AB50403	Catalog order: Order product three	382	0.609	0.601	0.000	6	4	5	5	5
AB91001	U.S. Savings Bonds	385	0.908	1.083	0.000	6	4	5	2	4
A110601	Registration & tuition Info	407	0.624	1.078	0.000	8	2	5	5	5
AB50301	Interest charges: Orally explain computation	433	0.602	1.523	0.000	2	5	5	5	7

13.4.3.5 Validity evidence for the quantitative scale

As with the prose tasks, one piece of validation evidence concerns the quantitative tasks newly developed for the 1990 survey. There were only 15 quantitative tasks in the 1985 young adult literacy assessment; to fill in the scale, 28 new tasks were developed for the 1990 assessment using the theoretical framework described here. As shown in Table 13-7, the combined set of structural complexity, process, and formulate variables accounts for 84 percent of the variance in scale values for the 28 new 1990 tasks and 83 percent of the variance for the combined set of 43 tasks from the combined task set from the 1985 and 1990 assessments.

Table 13-7. Unstandardized regression coefficients and standard errors predicting RP80 task difficulties on the basis of five structure, process, and formulate variables: 28 new quantitative literacy tasks and 43 total quantitative literacy tasks from the 1990 survey of job-seekers

	New Tasks			All Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural complexity	4.81	1.60	.01	4.11	1.59	.01
Process Variables						
Type of match	1.25	4.58	.79	.06	3.82	.99
Plausibility of distractors	20.44	3.52	.00	21.21	3.69	.00
Formulate Variables						
Type of calculation	11.56	3.07	.00	10.76	2.25	.00
Operation specificity	8.23	2.76	.01	9.57	2.44	.00
Variance accounted for:						
R^2		84%			83%	
Adjusted R^2		81%			81%	
Degrees of freedom		22			37	

13.4.4 Establishing Proficiency Levels

As the public report for the 1990 assessment of job-seekers was being drafted, the need to clarify the meaning of the scales as well as the graphic presentations of the results became increasingly clear. The challenge was to find a way to avoid the information overload of the figures and tables that were produced for the 1985 young adult literacy assessment (see Exhibit 13-2) and to pass along our growing understanding of appropriate score interpretation.

Empirical data of the kind presented in this chapter for each of the three literacy scales provides evidence for the credibility of the notion that while literacy is not a single skill suited to all types of tasks, neither is it an infinite number of skills each associated with a given prose, document, or quantitative task. Rather, as the data presented here indicate, there appears to be an ordered set of information-processing skills and strategies that may be called into play to accomplish the range of tasks represented by the three literacy domains.

As the project team sought effective ways to present the data from the 1990 survey, they re-examined the implications of the increases in process complexity as task difficulty rises. As tasks became more difficult, their associated code values on the task variables also increased (See Tables 13-2, 13-4, and 13-6). This relationship between task difficulty and code values appeared to be quite systematic. That is, toward the bottom of each literacy scale, the code value of 1 on each task variable was dominant; values of 2 and 3 became more frequent as tasks moved up the prose, document, and quantitative scales;

and toward the higher end, code values of 4, 5, and higher become predominant. Although the patterns differed somewhat from scale to scale reflecting differences in the coded values assigned to the variables, major shifts in the processes and skills required for successful task performance occurred at remarkably similar points on the three scales.

Visual inspection of the distributions of task codes along each of the literacy scales revealed several major points occurring at roughly 50 point intervals beginning with 225 on each scale. As with all systems, this one contains some noise and does not account for all of the score variance associated with performance on the three literacy scales. Moreover, the shifts in skill or process requirements do not necessarily occur at exactly 50 point intervals on the scales. However, assigning the exact range of scores to capture each level (for example, using score 277-319 to represent Level 3 on the document scale and 331-370 to represent Level 4 on the quantitative scale) implies a precision of measurement that is inappropriate for the methodology adopted. In order to ensure consistency across the scales, 50 point intervals were imposed. The advantage of having common intervals outweighs the marginal gain in agreement percentages and offers better readability to the users. Consequently, on the basis of the distributions of processing requirements, the results of the 1990 survey were aggregated into five proficiency levels depending on the range of the scores: Level 1 (less than or equal to 225), Level 2 (from 226 to 275), Level 3 (from 276 to 325), Level 4 (from 326 to 375), and Level 5 (greater than or equal to 376).

Once the levels were tentatively set, criteria to account for task placement within levels were determined, based solely on inspection of the code values assigned to each task. These criteria and the percentages of tasks meeting these criteria are shown in Table 13-8. Overall, an average of 78 percent of the prose tasks met the identified criteria for each level. The agreement percentages on the document and quantitative scales were 89 and 79, respectively. The advantage of having common intervals across scales outweighs the marginal gain in agreement percentages, thus were implemented.

Table 13-8. Numerical criteria assigned to task variables to distinguish proficiency levels and percentages of tasks meeting those criteria for prose, document, and quantitative literacy scales

	Prose literacy		Document literacy		Quantitative literacy	
	Criteria[*]	Percent agreement	Criteria[H]	Percent agreement	Criteria[I]	Percent agreement
Level 1	1, 1, ≤2	100	1, ≤2, ≤2	87	1, 1, ≤2	100
Level 2	2, 2, 2 or 3, ≤3, ≤3	73	2, 2, 2 or 3, ≤3, ≤2	97	3, ≤2, ≤2	50
Level 3	4, ≤3, ≤3	80	≥4, ≤3, ≤3	87	≤5, ≤3, ≤3	82
Level 4	4, ≥4, ≤4	73	4, ≥4, ≤3	75	≥4, ≤4, ≤4	77
Level 5	5, ≥5, ≤5	75	5, ≥5, ≤5	86	5, ≥5, ≤5	80
Overall agreement		78		89		79

[*] Criteria for prose literacy tasks pertain to type of match, plausibility of distractors, and type of information. [H] Criteria for document literacy tasks pertain to type of match, plausibility of distractors, and type of information. [I] Criteria for quantitative literacy tasks pertain to operation specificity, type of calculation, and plausibility of distractors.

The next step in evaluating the utility of using the five identified levels as reporting categories was to run regression analyses using levels rather than individual task scale values as the dependent variable. These results are presented in Table 13-9. As shown here, the identical process variables are significant in predicting proficiency levels as was the case in predicting specific task scale values. Moreover, the models used to predict proficiency levels account for roughly the same amount of variance as those used to predict task values—ranging from 78 percent on the quantitative scale to 80 percent on the prose scale to 88 percent on the document scale. These data are somewhat surprising given the typical effects of restriction of range on correlational data.

Table 13-9. Unstandardized regression coefficients and standard errors predicting five levels of RP80 task difficulty on the basis of six structure, process, and formulate variables: All prose, document, and quantitative literacy tasks from the 1990 survey of job-seekers

	Prose literacy			Document literacy			Quantitative literacy		
	Coeff	StdErr	p	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable									
Readability/structural complexity	.03	.04	.53	.00	.20	.93	.07	.37	.07
Process Variables									
Type of match	.19	.08	.03	.46	.04	.00	-.05	.09	.58
Plausibility of distractors	.49	.09	.00	.43	.04	.00	.48	.09	.00
Type of information	.25	.08	.01	.21	.06	.00	-	-	-
Formulate Variables									
Type of calculation	-	-	-	-	-	-	.14	.05	.01
Operation specificity	-	-	-	-	-	-	.20	.06	.00
Variance accounted for:									
R^2		80%			88%			78%	
Adjusted R^2		78%			87%			75%	

In addition to accounting for significant amounts of variance, the variables described in this chapter illustrate the internal consistency of processing characteristics associated within each of the five levels. This, in turn, provides evidence of the substantive aspect of construct validity, or the appropriateness of the theoretical model predicted to underlie consistencies in performance. Given this evidence concerning the validity of these five levels, it was decided to use them to report the distributions of the literacy of job-seeking adults participating in the two Department of Labor programs (the Job Training Partnership Act and the U.S. Employment Service) and to communicate the meaning of what was being measured along each of the scales.

Brief statements were developed to summarize some of the knowledge and skills associated with successful performance within each of the identified levels. These descriptions were derived from the pattern of codes among the processing variables associated with tasks falling within a level and are shown in Table 13-10. Displaying tasks along each scale as was done for the 1985 young adult literacy assessment (see Exhibit 13-36) tends to encourage interpretation of the scale at the task level. Rather than simply displaying tasks along each scale, the use of the five levels allowed the development of descriptive information that could be combined with the percentages of adults in various subpopulations who demonstrated performance within each of the designated levels.

Table 13-10. Descriptions of prose, document, and quantitative literacy levels based on type of match, plausibility of distractors, type of information, operation specificity, and type of calculation: 1990 survey of job-seekers

Levels	Prose literacy	Document literacy	Quantitative literacy
Level 1 0-225	Prose tasks at this level are the least demanding in terms of what the reader must do to produce a correct response. Typically, tasks at this level require the reader to locate one piece of information in which there is a literal match between the question and the stimulus material. If a distractor or plausible answer is present, it tends to be located away from where the correct information is found.	Tasks at this level are the least demanding. In general, they require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	Although no quantitative tasks used in this assessment fall within this level, experience suggests that such tasks would require a single, relatively simple operation for which the numbers are given and the operation specified.
Level 2 226-275	Some of the prose tasks of this level still require the reader to locate on a single literal feature of information; however, these tasks tend to occur in materials where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the readers to integrate information by pulling together two or more pieces of information or by comparing and contrasting information.	Tasks at this level begin to become more varied. Some still require the reader to match a single match a single piece of information; however, tasks occur where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the reader to cycle through information or to integrate information.	Tasks at this level typically require the use of a single operation based on numbers that are either stated in the question or easily located in the material. In addition, the operation needed is either stated in the question or easily determined based on the format of the problem—for example, entries on a bank deposit slip or order form.
Level 3 276-325	Tasks at this level tend to require the reader to search fairly dense text for literal or synonymous matches on the basis of more than one feature of information or to integrate information from relatively long text that does not contain organizational aids such as headings.	Tasks at this level tend to require the reader to either integrate three pieces of information or to cycle through materials in rather complex tables or graphs in which distractor information is present.	What appears to distinguish tasks at this level is that two or more numbers needed to solve the problem must be found in the stimulus material. Also the operation(s) needed can be determined from arithmetic relation terms.
Level 4 326-375	Tasks at this level continue to demand more from the reader. Not only are multiple-feature matching and integration of information from complex displays materials maintained, the degree of inferencing required by the reader is also increased. Conditional information is frequently present in tasks at this level that must be taken into account.	Tasks at this level tend to demand more from the reader. Not only are multiple-feature matching, cycling, and integration of information maintained, the degree of inferencing is increased. Cycling tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	Quantitative tasks at level 4 tend to require two or more sequential operations or the application of a single operation where either the quantities must be located in complex displays and/or the operation must be inferred from semantic information given or prior knowledge.
Level 5 376-500	These tasks require the reader to search for information in dense text or complex documents containing multiple plausible distractors, to make high text-based inferences or use specialized background knowledge, as well as to compare and contrast sometimes complex information to determine differences.	Tasks at this level require the most from the reader. The reader must search through complex displays contain[ing] multiple distractors, make high text-based inferences or use specialized knowledge.	Quantitative tasks at this level are the most demanding. They tend to require the reader to perform multiple operations and to disembed features of a problem from stimulus material or to rely on background knowledge to determine the quantities or operations needed.

In addition, using information derived from the IRT analyses, it became possible to estimate the likelihood that individuals with various proficiency levels would perform the average task within a specified level correctly. These distributions of likelihood estimates provide a richer and more accurate reflection of the range of tasks that an individual can be expected to perform successfully. Collectively, the descriptors, distributions, and probabilities of correctly performing various tasks within different levels could be displayed in a single table. An example, taken from the 1990 public report (Kirsch, Jungeblut and Campbell, 1992), is reproduced as Exhibit 13-36.

BEST COPY AVAILABLE

Exhibit 13-36. Descriptions of five document literacy levels, average RP80 scale values, and probabilities of performing tasks in the level successfully: Adult job seekers, 1990

Levels	Description of Document Tasks at Each of Five Levels	Average RP 80 at Each Level	Average Probability at Selected Proficiency Levels						TOTAL
			200	250	300	350	400		
Level 1 01-225	Tasks at this level are the least demanding. In general, they require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	194	80	94	98	100	100	JTPA	14.1 (2.0)
								ES/UI	13.1 (1.6)
								Young Adults	8.0 (0.6)
Level 2 226-275	Tasks at this level begin to become more varied. Some still require the reader to math a single piece of information; however, tasks occur where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the reader to cycle through information or to integrate information.	248	47	81	95	99	100	JTPA	37.3(1.3)
								ES/UI	30.1 (1.2)
								Young Adults	24.2 (1.1)
Level 3 276-325	Tasks at this level tend to require the reader to either integrate three pieces of information or to cycle through materials in rather complex tables or graphs in which distractor information is present.	300	30	54	79	93	97	JTPA	35.4 (1.5)
								ES/UI	35.9 (1.0)
								Young Adults	39.7(1.2)
Level 4 326-375	Task at this level continue to demand more from the reader. Not only are multiple-feature matching, cycling, and integration of information maintained, the degree of inferencing is increased. Cycling tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	351	11	26	53	79	93	JTPA	12.2 (1.8)
								ES/UI	18.5 (1.7)
								Young Adults	24.0 (1.1)
Level 5 376-500	Tasks at this level require the most from the reader. The reader must search through complex displays containing multiple distractors, make high text-based inferences, or use specialized knowledge.	405	15	23	37	60	79	JTPA	1.1 (0.4)
								ES/UI	2.4 (0.5)
								Young Adults	4.1 (0.6)

*The numbers in parentheses are estimated standard errors.

It will be seen, for example, that about 37 percent of the JTPA participants demonstrated performance in the Level 2 range of the document scale. The average difficulty of tasks in this level is 248, and an individual scoring at 250 has a probability of 81 percent of performing such a task successfully. Individuals scoring at 300 and above are likely to make few errors on tasks at around 248 on the document scale. Similarly, an individual with a proficiency score of 250 has a better than 90 percent chance of responding correctly to tasks in Level 1. This same individual has a probability of about 50 percent of successfully performing Level 3 tasks and about a 25 percent probability of performing Level 4 and Level 5 tasks correctly.

13.5 THE 1992 NATIONAL ADULT LITERACY SURVEY

The 1992 National Adult Literacy Survey assessed the literacy skills of a nationally representative sample of individuals age 16 and older, as well as representative samples of individuals ages 16-64, in 12 states (Kirsch, Jungeblut, Jenkins, and Kolstad, 1993). The National Center for Education Statistics requested

BEST COPY AVAILABLE

that the assessment results be linked to both the 1985 young adult literacy assessment and the 1990 survey of adult job-seekers served by the U.S. Department of Labor. To help meet the survey's objectives, a Literacy Definition and a Technical Review Committee were established. After some deliberations, members of the Literacy Definition Committee recommended the adoption of the same definition and measurement framework used in the two earlier surveys. In addition, members of this committee also requested that the results from the 1992 National Adult Literacy Survey be reported in terms of the same five levels developed for the 1990 survey of job-seekers. This decision provided a unique opportunity to investigate further the validity of the theoretical framework that had evolved for developing new assessment tasks and for enhancing score meaning.

13.5.1 Prose Literacy

The development of the 1992 National Adult Literacy Survey entailed both the reuse of existing prose literacy tasks from the 1985 young adult literacy assessment and the production of new prose literacy tasks. The new tasks continued the emphasis on expository prose drawn from authentic sources that adults might ordinarily encounter in daily life. The resulting assessment pool for the 1992 National Adult Literacy Survey included 41 prose literacy tasks of which 27 were newly developed for the 1992 survey.

A preceding section described in detail the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the prose scale. Just as with the 1990 survey and the 1985 young adult literacy assessment, these rules were applied to all prose literacy tasks in the 1992 National Adult Literacy Survey, and the resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-11.

Table 13-11. List of new prose literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: The 1992 National Adult Literacy Survey

Identifier	Task Description	Scaled RP80	IRT parameters			Read-ability	Type of match	Distractor Plausibility	Information type
			a	b	c				
NC00301	"My Dream:" Find country in short story	150	0.893	-3.228	0.000	4	1	1	1
N120901	Susan Butcher: Find number of wins of sled race	210	0.889	-2.061	0.000	9	1	1	2
NC00401	"My Dream:" Underline sentence explaining action	224	0.765	-1.936	0.000	4	1	2	4
N080101	SSI: Mark correct information in article	226	1.329	-1.447	0.000	6	1	1	3
N100101	"Growing Up:" Find first buyer's name	239	1.467	-1.147	0.000	8	3	2	1
N090601	Face off: What group will mandate safe cars?	253	1.878	-0.748	0.000	10	3	2	1
N090701	Face off: Find correct information in article	256	1.805	-0.699	0.000	10	3	2	2
N110101	Blood pressure: Why difficult to know if high	262	0.988	-0.971	0.000	7	3	2	4
N130801	Cost to raise child: Find information from article	274	0.735	-1.013	0.000	6	2	4	2
N110501	Jury: Underline sentence explaining action	276	0.939	-0.731	0.000	7	4	3	3
N080201	SSI: What must an SSI user accept if offered?	277	1.516	-0.389	0.000	6	4	2	3
N100201	"Growing Up:" Determine correct day of delivery	284	1.297	-0.346	0.000	8	4	3	2
N100301	"Growing Up:" What reason given to stop selling?	287	1.187	-0.344	0.000	8	5	1	4
N010201	Marketing: Underline sentence explaining action	288	1.059	-0.403	0.000	15	3	4	3
N110401	Jury: Length of time served by a juror	314	0.770	-0.192	0.000	7	4	2	4
N120301	Ida Chen: What experience turned Ida toward law?	316	1.075	0.142	0.000	7	4	2	3
N120401	Two things Chen did to resolve discrimination conflicts	317	1.162	0.229	0.000	7	4	3	2
N130201	"Fueled:" Determine phrase meaning	324	1.089	0.316	0.000	9	5	1	3
N130401	"Fueled:" Give suggestion about good value change	346	1.576	0.979	0.000	9	5	1	4
N010101	Marketing: List two facts	349	0.869	0.608	0.000	15	5	5	4
N090801	Contrast views on fuel-efficiency vs. Size of car	360	1.239	1.091	0.000	10	6	2	5
N080301	SSI: What is most you can make to receive SSI?	362	0.619	0.486	0.000	6	4	5	2
N130301	"Fueled:" Give diff and similarity between events	375	0.978	1.214	0.000	9	6	2	4
N100401	"Growing Up:" Compare approaches to selling mags	383	0.842	1.236	0.000	8	6	2	5
N110601	Two challenges attorneys use to jurors	410	1.045	1.954	0.000	6	6	2	5
N120501	Ida Chen: Interpret phrase from article	424	0.927	2.107	0.000	7	6	3	5
N010301	Marketing: Give purpose of event	433	0.787	2.138	0.000	15	5	5	3

Another piece of validation evidence can be obtained from analysis of the tasks used in the 1992 National Adult Literacy Survey. The four task variables were used in a regression analysis designed to predict the RP80 scale values of each task. Table 13-12 shows the results of these analyses for both the new tasks as well as the entire 1992 pool of prose literacy tasks. The numbers in the tables represent the raw beta coefficients for each of the variables included in the regression analyses along with standard errors and probabilities. Overall, the three process variables were significant for both the new 1992 tasks and for the complete set of prose tasks. The same variables as were found with the 1990 prose literacy tasks contribute to the predictive models. In addition, the amount of variance accounted for in the new tasks is similar—81 percent on the 1990 survey and 89 percent on the 1992 survey.

Table 13-12. Unstandardized regression coefficients and standard errors predicting prose literacy RP80 task difficulties on the basis of four structure and process variables: 27 new prose literacy tasks and 41 total prose literacy tasks from the 1992 survey of adults.

Prose literacy	New 1992 Tasks			All 1992 Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Readability	-.07	1.80	.68	.18	1.67	.69
Process Variables						
Type of match	29.65	3.62	.00	28.86	3.45	.00
Plausibility of distractors	18.41	4.06	.00	16.09	3.63	.00
Type of information	10.91	4.35	.02	8.84	4.17	.04
Variance accounted for:						
R^2		89%			87%	
Adjusted R^2		87%			86%	
Degrees of freedom		22			36	

Since the equations predicting prose literacy task difficulty in the 1992 data essentially reproduced the findings of the 1990 data, and since there was a contractual requirement for comparability with the prior surveys, there was no need to revisit the cutpoints for the literacy levels. Still, the 1992 data provided an occasion to improve the language describing the literacy levels in minor ways. Table 13-13 provides a comparison of verbal descriptions of the prose literacy levels used in reporting the 1990 and 1992 survey results. The minor adjustments that can be seen in the 1992 descriptions were designed to make them more consistent with the variables predicting task difficulty.

Table 13-13. Descriptions of prose literacy levels based on type of match, plausibility of distractors, and type of information: 1990 survey of job-seekers and the 1992 National Adult Literacy Survey.

Levels	1990 Description	1992 Description
Level 1 0-225	Prose tasks at this level are the least demanding in terms of what the reader must do to produce a correct response. Typically, tasks at this level require the reader to locate one piece of information in which there is a literal match between the question and the stimulus material. If a distractor or plausible answer is present, it tends to be located away from where the correct information is found.	Most of the tasks in this level require the reader to read relatively short text to locate a single piece of information which is identical to or synonymous with the information given in the question or directive. If plausible but incorrect information is present in the text, it tends not to be located near the correct information.
Level 2 226-275	Some of the prose tasks of this level still require the reader to locate on a single literal feature of information; however, these tasks tend to occur in materials where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the readers to integrate information by pulling together two or more pieces of information or by comparing and contrasting information.	Some tasks in this level require readers to locate a single piece of information in the text; however, several distractors or plausible but incorrect pieces of information may be present, or low-level inferences may be required. Other tasks require the reader to integrate two or more pieces of information or to compare and contrast easily identifiable information based on a criterion provided in the question or directive.
Level 3 276-325	Tasks at this level tend to require the reader to search fairly dense text for literal or synonymous matches on the basis of more than one feature of information or to integrate information from relatively long text that does not contain organizational aids such as headings.	Tasks in this level tend to require readers to make literal or synonymous matches between the text and the information given in the task, or to make matches that require low-level inferences. Other tasks ask readers to integrate information from dense or lengthy text that contains no organizational aids such as headings. Readers may be asked to generate a response based on information that can be easily identified in the text. Distracting information is present, but is not located near the correct information.
Level 4 326-375	Tasks at this level continue to demand more from the reader. Not only are multiple-feature matching and integration of information from complex displays materials maintained, the degree of inferencing required by the reader is also increased. Conditional information is frequently present in tasks at this level that must be taken into account.	These tasks require readers to perform multiple-feature matches and to integrate or synthesize information from complex or lengthy passages. More complex inferences are needed to perform successfully. Conditional information is frequently present in tasks at this level and must be taken into consideration by the reader.
Level 5 376-500	These tasks require the reader to search for information in dense text or complex documents containing multiple plausible distractors, to make high text-based inferences or use specialized background knowledge, as well as to compare and contrast sometimes complex information to determine differences.	Some tasks in this level require the reader to search for information in dense text which contains a number of plausible distractors. Others ask readers to make high-level inferences or use specialized background knowledge. Some tasks ask readers to contrast complex information.

13.5.2 Document Literacy

Just as with the prose literacy scale, the development of the 1992 National Adult Literacy Survey entailed both the reuse of existing document literacy tasks from the 1985 young adult literacy assessment and the production of new document literacy tasks. The new tasks continued the emphasis on documents drawn from authentic sources that adults might ordinarily encounter in daily life. The resulting assessment pool for the 1992 National Adult Literacy Survey included 82 document literacy tasks of which 26 were newly developed for the 1992 survey.

Previous sections of this chapter have described in detail and illustrated with examples the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the document literacy scale. Just as with the 1990 survey and the 1985 young adult literacy assessment, these rules were applied to all document literacy tasks in the 1992 National Adult Literacy Survey. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-14.

Table 13-14. List of new document literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: The 1992 National Adult Literacy Survey

Identifier	Task Description	Scaled RP80	IRT parameters			Complexity	Type of match	Distractor Plausibility	Information type
			a	b	c				
SCOR100	Social Security card: Sign name on line	70	0.505	-4.804	0.000	1	1	1	1
SCOR300	Driver's license: Locate expiration date	152	0.918	-2.525	0.000	2	1	2	1
N090301	Essence: Determine page certain article begins on	211	1.124	-1.225	0.000	3	1	2	2
N120601	Middle class: Find projected percent	213	0.795	-1.488	0.000	4	3	2	2
N010401	Vehicle chart: Find correct information	215	0.903	-1.341	0.000	2	1	2	2
N080802	Auto maintenance form: Enter given information	233	1.357	-0.684	0.000	3	2	3	2
N120101	Campus map: Mark map for given info	239	0.986	-0.802	0.000	7	3	2	1
N130103	S.S. card application: Identify and enter info(2)	243	2.106	-0.291	0.000	5	2	3	1
N130102	S.S. card application: Identify and enter info(3)	243	1.270	-0.544	0.000	5	2	2	3
N110302	Certified mail rec't: Enter postage and fee	244	0.714	-1.026	0.000	2	2	2	2
N110301	Certified mail rec't: Enter name and address	251	0.812	-0.743	0.000	2	2	2	2
N130104	S.S. card application: Identify and enter info(4)	251	2.159	-0.111	0.000	5	2	2	1
N090401	Essence: Determine topic of given article	257	0.988	-0.448	0.000	3	2	2	3
N130101	S.S. card application: Identify and enter info(1)	259	1.619	-0.096	0.000	5	2	2	2
N080701	Bus schedule: Mark map correctly for given info	260	1.095	-0.312	0.000	9	3	2	1
N010801	Trend chart: Mark information on chart	266	0.808	-0.463	0.000	3	3	2	1
N120201	Campus map: Find correct room for given dean	267	0.842	-0.403	0.000	7	3	2	2
N090501	Essence: Determine topic of section of magazine	285	0.671	-0.301	0.000	3	4	2	3
N100501	Opinions table: Mark sentence explaining action	304	1.039	0.486	0.000	4	3	2	4
N080601	Bus schedule: Take correct bus for given condition	305	1.040	0.505	0.000	9	4	2	2
N011001	Trend chart: Determine least # of points needed	317	0.646	0.261	0.000	3	5	3	2
N080801	Auto maintenance form: Enter information given (1)	323	0.763	0.570	0.000	3	3	4	2
N110701	Credit card table: Find correct bank	335	0.470	0.126	0.000	5	4	4	1
N100601	Opinions table: Find correct group for given info	343	1.135	1.285	0.000	4	2	4	2
N010901	Trend chart: Put information on chart	386	0.721	1.702	0.000	3	5	5	5
N110901	Credit card table: Give 2 differences	388	0.829	1.883	0.000	5	8	2	5
N100701	Summarize views of parents & teachers	396	1.128	2.300	0.000	4	8	3	5

Just as with the prose literacy tasks, parallel analysis of the document literacy tasks used in the 1992 National Adult Literacy Survey provides another piece of validation evidence. The four task variables were again used in a regression analysis designed to predict the RP80 scale values of each task. Table 13-15 shows the results of these analyses for both the new document literacy tasks as well as the entire 1992 pool of document literacy tasks. Overall, the three process variables were significant for both the new 1992 tasks and for the complete set of document literacy tasks. Just as on the prose scale, the same variables contribute to the predictive models, and the amount of variance accounted for is similar—92 percent on the 1990 survey and 88 percent on the 1992 survey.

Table 13-15. Unstandardized regression coefficients and standard errors predicting document literacy RP80 task difficulties on the basis of four structure and process variables: 25 new document literacy tasks and 81 total document literacy tasks from the 1992 survey of adults

Document literacy	New 1992 Tasks			All 1992 Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural Complexity	2.83	2.58	.29	.09	1.07	.42
Process Variables						
Type of match	17.09	3.57	.00	20.33	2.27	.00
Plausibility of distractors	28.15	5.30	.00	22.19	2.74	.00
Type of information	9.84	5.07	.07	11.69	3.05	.00
Variance accounted for:						
R^2		88%			84%	
Adjusted R^2		85%			83%	
Degrees of freedom		20			76	

*Scoring was excluded from the regression equation due to the fact that it is an outlier.

Just as with the prose literacy scale, the equations predicting document literacy task difficulty in the 1992 data essentially reproduced the findings of the 1990 data. Again, there was no need to revisit the cutpoints for the document literacy levels, yet the language describing the literacy levels was improved in minor ways. Table 13-16 provides a comparison of the verbal descriptions of document literacy levels used in reporting the 1990 and 1992 survey results. The minor adjustments that can be seen in the 1992 descriptions were intended to clarify and systematize the language, but not to indicate any substantive changes in their meaning.

Table 13-16. Descriptions of document literacy levels based on type of match, plausibility of distractors, and type of information: 1990 survey of job-seekers and the 1992 National Adult Literacy Survey

Levels	1990 Description	1992 Description
Level 1 0-225	Tasks at this level are the least demanding. In general, they require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	Tasks in this level tend to require the reader either to locate a piece of information based on a literal match or to enter information from personal knowledge onto a document. Little, if any, distracting information is present.
Level 2 226-275	Tasks at this level begin to become more varied. Some still require the reader to match a single match a single piece of information; however, tasks occur where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the reader to cycle through information or to integrate information.	Some tasks in this level require the reader to match a single piece of information; however, several distractors may be present, or the match may require low-level inferences. Tasks in this level may also ask the reader to cycle through information in a document or to integrate information from various parts of a document.
Level 3 276-325	Tasks at this level tend to require the reader to either integrate three pieces of information or to cycle through materials in rather complex tables or graphs in which distractor information is present.	Some tasks in this level require the reader to integrate multiple pieces of information from one or more documents. Others ask readers to cycle through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.
Level 4 326-375	Tasks at this level tend to demand more from the reader. Not only are multiple-feature matching, cycling, and integration of information maintained, the degree of inferencing is increased. Cycling tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	Tasks in this level, like those at the previous levels, ask readers to perform multiple-feature matches, cycle through documents, and integrate information; however, they require a greater degree of inferencing. Many of these tasks require readers to provide numerous responses but do not designate how many responses are needed. Conditional information is also present in the document tasks at this level and must be taken into account by the reader.
Level 5 376-500	Tasks at this level require the most from the reader. The reader must search through complex displays contain[ing] multiple distractors, make high text-based inferences or use specialized knowledge.	Tasks in this level require the reader to search through complex displays that contain multiple distractors, to make high-level, text-based inferences, and to use specialized knowledge.

13.5.3 Quantitative Literacy

As with the prose and document literacy scale, the development of the 1992 National Adult Literacy Survey entailed both the reuse of existing quantitative literacy tasks from the 1985 young adult literacy assessment and the production of new quantitative literacy tasks. The new tasks continued the emphasis on prose texts and documents drawn from authentic sources that adults might ordinarily encounter in daily life. The resulting assessment pool for the 1992 National Adult Literacy Survey included 43 quantitative literacy tasks of which 28 were newly developed for the 1992 survey.

Previous sections of this chapter have described in detail and illustrated with examples the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the quantitative literacy scale. Just as with the 1990 survey and the 1985 young adult literacy assessment, these rules were applied to all quantitative literacy tasks in the 1992 National Adult Literacy Survey. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-17.

Table 13-17. List of quantitative literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: The 1992 National Adult Literacy Survey

Identifier	Quantitative Literacy Items	Scaled RP80	IRT parameters			Complexity	Type of match	Distractor Plausibility	Calculation type	Operation specificity
			a	b	c					
NC00501	Checking: Total bank deposit entry	192	0.661	-2.792	0.000	2	1	1	1	1
N110303	Certified mail rec't: postage and fees	239	0.790	-1.730	0.000	3	1	2	1	2
NC00601	Price diff: Sleuth & On the Town	247	0.718	-1.690	0.000	2	2	1	2	2
N100801	Salt River: Determine diff in costs	251	0.648	-1.738	0.000	5	2	2	2	3
N101001	Salt River: hours between points	278	0.944	-0.838	0.000	5	3	4	1	3
N090901	Carpet ad: Diff in reg and sale price	278	0.790	-1.004	0.000	2	3	1	2	4
N120701	Pct diff black & white middle class	280	0.909	-0.845	0.000	4	4	2	2	4
N090201	Get net total owed after deduction	284	1.677	-0.349	0.000	3	2	4	2	2
N010501	Vehicle chart: sum of percentages	287	0.851	-0.769	0.000	2	2	2	1	2
N110801	Credit card table: Difference in rates	300	0.882	-0.495	0.000	4	1	2	2	4
N130601	Rec room: num wall panels needed	307	1.112	-0.184	0.000	5	1	2	4	3
N090101	Discount if oil bill paid in 10 days	309	1.347	-0.018	0.000	3	2	2	3	5
N080501	Time: student union to 17th & Main	322	0.757	-0.248	0.000	9	4	2	2	5
N011101	Gas gauge: show calculations	330	1.035	0.196	0.000	2	3	1	5	6
N121001	Miles/day Butcher went	332	1.018	0.218	0.000	2	3	2	4	6
N100901	Salt River: Miles between stops	334	0.623	-0.264	0.000	5	2	4	1	5
N010701	Vehicle chart: magnitude of diff	342	1.034	0.411	0.000	2	3	4	4	3
N081001	Rank juices by expense, w/ reasons	344	0.733	0.122	0.000	2	4	1	3	6
N130901	Money needed to raise child	351	0.946	0.499	0.000	2	1	2	3	5
N010601	Vehicle chart: Solution to pct problm	355	1.122	0.717	0.000	2	1	2	3	6
N110201	Blood pressure: Death rate	360	1.033	0.741	0.000	2	1	2	3	6
N120801	Middle class: Diff in size of pct	366	1.013	0.831	0.000	4	3	2	4	4
N080401	Yrly amount for couple w/ basic SSI	369	0.696	0.521	0.000	2	2	4	2	6
N080901	Auto form: Calculate miles/gallon	376	0.851	0.856	0.000	3	3	2	5	4
N130501	RecRoom: Feet of molding needed	389	0.655	0.819	0.000	5	2	3	5	7
N121101	Butcher: diff in completion times	406	0.960	1.518	0.000	2	2	1	5	5
N091001	Carpet ad:Total cost to carpet room	421	0.635	1.371	0.000	2	1	2	5	7
N130701	Rec room: Describe solution	436	0.846	1.962	0.000	5	2	2	5	7

*Structural Complexity

As with the prose and document literacy tasks, parallel analysis of the quantitative literacy tasks used in the 1992 National Adult Literacy Survey provides more evidence of construct validity. The four task variables were again used in a regression analysis designed to predict the quantitative literacy RP80 scale values of each task. Table 13-18 shows the results of these analyses for both the new quantitative literacy tasks as well as the entire 1992 pool of quantitative literacy tasks. Overall, the task variables were significant for both the new 1992 tasks and for the complete set of quantitative literacy tasks. Just as on the prose and document literacy scales, the same variables contribute to the predictive models, and the amount of variance accounted for is similar—84 percent on the 1990 survey (Table 13-7) and 84 percent on the 1992 survey (Table 13-18). Moreover, the similarity of results between the pools of the 1990 and 1992 tasks is especially important as evidence in support of the appropriateness of interpreting the 1992 survey results in terms of the five levels of literacy proficiencies established for reporting the 1990 survey results.

Table 13-18. Unstandardized regression coefficients and standard errors predicting quantitative literacy RP80 task difficulties on the basis of five structure, process, and formulate variables: 28 new quantitative literacy tasks and 43 total quantitative literacy tasks from the 1992 survey of adults

Quantitative literacy	New 1992 Tasks			All 1992 Tasks		
	Coeff	*StdErr*	*p*	Coeff	StdErr	*p*
Structure Variable						
Structural Complexity	.33	3.18	.92	.78	2.81	.78
Process Variables						
Type of match	-2.57	5.16	.62	.68	4.24	.87
Plausibility of distractors	12.25	5.16	.03	9.66	4.17	.03
Formulate Variables						
Type of calculation	20.75	4.62	.00	14.43	3.21	.00
Operation specificity	16.64	3.84	.00	18.39	2.94	.00
Variance accounted for:						
R^2		84%			81%	
Adjusted R^2		81%			78%	
Degrees of freedom		22			37	

As with the prose and document literacy scales, the equations predicting quantitative literacy task difficulty in the 1992 data essentially reproduced the findings of the 1990 data. Again, there was no need to revisit the cutpoints for the quantitative literacy levels, yet the language describing the literacy levels was adjusted in minor ways without affecting the substantive meaning of the descriptions. Table 13-19 provides a comparison of the verbal descriptions of quantitative literacy levels used in reporting the 1990 and 1992 survey results.

Table 13-19. Descriptions of quantitative literacy levels based on type of operation, plausibility of distractors, and type of calculation: 1990 survey of job-seekers and the 1992 National Adult Literacy Survey

Levels	1990 Description	1992 Description
Level 1 0-225	Although no quantitative tasks used in this assessment fall within this level, experience suggests that such tasks would require a single, relatively simple operation for which the numbers are given and the operation specified.	Tasks in this level require readers to perform single, relatively simple arithmetic operations, such as addition. The numbers to be used are provided and the arithmetic operation to be performed is specified.
Level 2 226-275	Tasks at this level typically require the use of a single operation based on numbers that are either stated in the question or easily located in the material. In addition, the operation needed is either stated in the question or easily determined based on the format of the problem—for example, entries on a bank deposit slip or order form.	Tasks in this level typically require readers to perform a single operation using numbers that are either stated in the task or easily located in the material. The operation to be performed may be stated in the question or easily determined from the format of the material (for example, an order form).
Level 3 276-325	What appears to distinguish tasks at this level is that two or more numbers needed to solve the problem must be found in the stimulus material. Also the operation(s) needed can be determined from arithmetic relation terms.	In tasks in this level, two or more numbers are typically needed to solve the problem, and these must be found in the material. The operation(s) needed can be determined from the arithmetic relation terms used in the question or directive.
Level 4 326-375	Quantitative tasks at level 4 tend to require two or more sequential operations or the application of a single operation where either the quantities must be located in complex displays and/or the operation must be inferred from semantic information given or prior knowledge.	These tasks tend to require readers to perform two or more sequential operations or a single operation in which the quantities are found in different types of displays, or the operations must be inferred from semantic information given or drawn from prior knowledge.
Level 5 376-500	Quantitative tasks at this level are the most demanding. They tend to require the reader to perform multiple operations and to disembed features of a problem from stimulus material or to rely on background knowledge to determine the quantities or operations needed.	These tasks require readers to perform multiple operations sequentially. They must disembed the features of the problem from text or rely on background knowledge to determine the quantities or operations needed.

13.5.4 Levels of Prose, Document, and Quantitative Literacy

In addition to the above analyses of continuous RP80 task difficulties, it was also important to analyze the 1992 survey tasks in terms of categorical literacy levels to make sure that the variables and amounts of explained variance were roughly the same. The results of these regression analyses are shown in table 13-20; the R^2s are in the range from 79 to 88 percent. These results are quite comparable with those reported in the 1985 and 1990 surveys. In addition, the amount of variance accounted for is nearly identical for the level regressions compared to the RP80 difficulty regressions—89 percent for continuous tasks compared to 88 percent for levels on the prose scale; 88 percent compared to 86 percent respectively on the document scale; and 81 percent compared to 79 percent on the quantitative scale.

Table 13-20. Unstandardized regression coefficients and standard errors predicting five levels of RP80 task difficulty on the basis of six structure and process variables: All prose, document, and quantitative literacy tasks from the 1992 survey of adults

	Prose literacy			Document literacy			Quantitative literacy		
	Coeff	StdErr	p	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable									
Structural Complexity	.01	.03	.65	.00	.20	.93	.07	.37	.07
Process Variables									
Type of match	.53	.06	.00	.42	.04	.00	.04	.08	.64
Plausibility of distractors	.26	.06	.00	.43	.05	.00	.21	.79	.01
Type of information	.15	.07	.05	.23	.05	.00	-	-	-
Formulate Variables									
Type of calculation	-	-	-	-	-	-	.21	.06	.00
Operation specificity	-	-	-	-	-	-	.35	.06	.00
Variance accounted for:									
R^2	88%			86%			79%		
Adjusted R^2	86%			86%			76%		

13.6 CONCLUSION

One of the goals of large-scale surveys is to provide a set of information that can inform the decision-making process. Important to this goal is presenting data in a manner that will enhance the understanding of what has been measured and of the conclusions that may be drawn both within and across assessments. The theoretical model that has evolved through three literacy assessments using the same definition and measurement framework has been a useful and valid way to report on the condition of adult literacy in America. This model identifies a set of variables that has been shown to underlie successful performance on a broad array of literacy tasks across several surveys. These variables, in turn, have been useful in developing new literacy tasks that help us to refine and extend our measurement of literacy. Moreover, they provide a framework for understanding what is being measured that allows us to identify levels of performance that have generalizability and validity across assessments and groups, rather than interpreting results in terms of discrete tasks.

Collectively, the knowledge and understanding that derives from such models contributes to an evolving conception of test design that begins to move away from merely assigning a numerical value (or position) to an individual based on responses to a set of tasks and toward assigning meaning and interpretability to this number. This, in turn, provides evidence of the appropriateness of the theoretical models predicted to underlie consistencies in performance.

Chapter 14

LITERACY LEVELS AND THE 80 PERCENT RESPONSE PROBABILITY CONVENTION

Andrew Kolstad, National Center for Education Statistics

As described in the preceding chapter of this report, the 1992 National Adult Literacy Survey employed a response probability convention of 80 percent to locate literacy tasks on the three literacy scales, and then used the locations of the tasks to set boundaries between the literacy levels. Many users of this survey do not realize how sensitive the basic survey findings are to small variations in the response probability convention. If an alternative response probability convention were to be used to locate literacy tasks on the scale, the boundaries between the literacy levels would shift.

The initial survey findings reported that large proportions of adults perform in Levels 1 and 2 on the three literacy scales, and that such adults were at risk due to their limited literacy skills. Nevertheless, most adults in Levels 1 and 2 reported that they could read and write English "well" or "very well." Stitch and Armstrong (1994) were the first to point out that the 80 percent response probability criterion used in the adult literacy survey could be a possible source of the reported gap between performance and perception.

This chapter explores the importance of the response probability convention in reporting prose literacy results from the 1992 National Adult Literacy Survey. The chapter will show that raising or lowering the response probability convention would distribute the population differently across the five levels of literacy. However, such changes would not mean that adults have more literacy skills than previously reported. The true literacy proficiencies of the population remain as reported. What would change is the line that separates those who are classified as "able to do" the literacy tasks and those who are not. The response probability convention influences how the results are interpreted, not how well adults perform on the assessment.

The following sections of this chapter describe the prose literacy tasks and their characteristics; the use of item mapping to anchor the prose literacy scale by locating specific tasks along it (using a response probability convention); the need for a response probability convention; the relationship of the response probability convention to the cut points between the literacy levels; and the variation in the proportions of the adult population reported to be in each prose literacy level as a function of the response probability convention. The final part of the paper discusses a few implications of the findings.

14.1 PROSE LITERACY TASKS AND THEIR CHARACTERISTICS

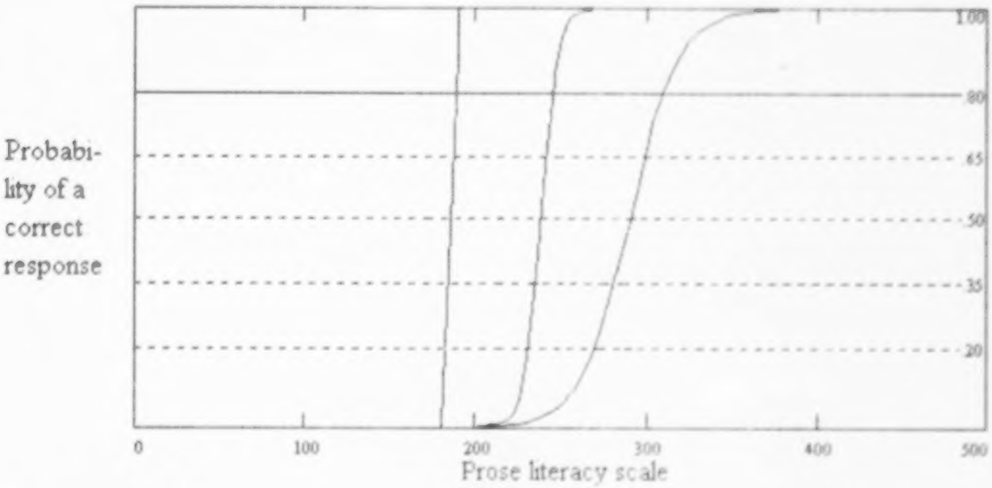
The results of this investigation are based on survey responses and assessment data from the 1992 National Adult Literacy Survey, supplemented with task-specific data developed by Mosenthal and Kirsch that served as the basis of the literacy levels (see Chapter 13). The 1992 National Adult Literacy Survey included 165 literacy tasks, of which 41 were used to measure prose literacy. A typical adult in the survey responded to 11 prose literacy tasks—not enough to measure any particular adult's prose literacy skills with any accuracy, but with a large sample, the 41 litaracy tasks are enough to estimate the distribution of prose literacy skills in the adult population. Only two of the prose literacy tasks were multiple-choice items, while the remainder were answered with a word, phrase, or a few short sentences. All were scored as right or wrong, with no partial credit allowed.

The assessment tasks in the survey were designed to measure prose literacy as a unidimensional scale. Ideally, one ought to be able to predict with a good deal of accuracy the pattern of right and wrong answers to the prose literacy tasks in the assessment from a person's score on the prose literacy scale. Success with prose tasks would also be cumulative. Once any particular skill is mastered, any task needing that skill could be performed correctly. As described in Chapter 9, the 1992 National Adult Literacy Survey used item response theory (IRT) to model the relationship between proficiency—an unobservable variable that is estimated from the responses to many test questions—and correct responses to any particular test question (Lord and Novick, 1968). A feature of such models that provides some realism is that the likelihood of a correct response does not jump immediately from zero to one at some point along the proficiency scale, but rises more gradually as a function of proficiency.

If a real literacy task had a high enough discrimination parameter, its item characteristic curve (ICC) would look like a step-function. Test developers try to create tasks that vary in their difficulty and have the highest possible discrimination parameters, but they do not achieve this kind of perfection. Figure 14-1 displays three item characteristic curves representing hypothetical literacy tasks. Here, the vertical axis represents the probability of a correct response, while the horizontal axis represents the score on the 0-to-500 prose literacy scale. The 1992 National Adult Literacy Survey associated each literacy task with the point on the scale at which an adult would have an 80 percent chance of success with that task. The other horizontal guidelines in the figure (equal to .20, .35, .50, and .65) are presented for comparative purposes, as discussed below. What this figure reveals is that as proficiency increases, so does the likelihood of a correct response to each item.

In Figure 14-1, the ICC of the literacy task on the left was generated by a hypothetical IRT function with its discrimination parameter set high enough to approximate a step function ($a = 50$), while the one on the right was generated by a hypothetical IRT function with a discrimination parameter set equal to the highest value that actually occurred among any of the 165 literacy tasks ($a = 2.16$). With the task on the left, the boundary between success and failure is very narrow. The probability of success changes from zero to one within a range of about 5 points on the literacy scale. With the task on the right, the boundary between success and failure on this task is more broad. Anyone with a prose literacy score in the range between roughly 250 and 350 has a probability of success that (after rounding) is neither zero nor one. In this range, success or failure with such a task is not a certainty. Yet this task discriminated better than any other.

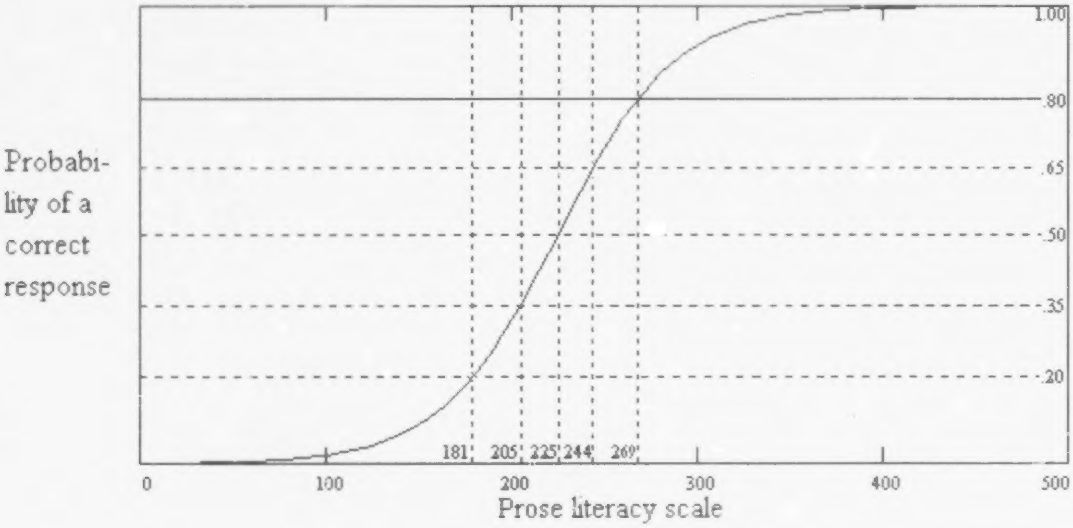
Figure 14-1. Three hypothetical tasks of varying discrimination (and difficulty)



An average prose literacy task from the 1992 National Adult Literacy Survey did not discriminate as well as the best task shown in Figure 14-1. To illustrate performance on an average prose literacy task, Figure 14-2 below presents a hypothetical ICC with average values of both discrimination and difficulty, based on the tasks actually used in the 1992 National Adult Literacy Survey. The intersections between the horizontal guidelines (at probabilities of .20, .35, .50, .65, and .80) and the item characteristic curve locate points on the prose literacy scale. These are the points at which prose proficiency is sufficient to achieve these probabilities of success on this hypothetical task.

What Figure 14-2 reveals is that as proficiency increases, so does the likelihood of a correct response. Typically this increase is gradual. The 1992 National Adult Literacy Survey associated each literacy task with the point on the scale at which an adult would have an 80 percent chance of success with that task. For an average prose task, a proficiency score of at least 269 would be needed to predict success with the 80 percent response probability adopted as a standard by the 1992 National Adult

Figure 14–2. Hypothetical prose literacy task with average difficulty and discrimination

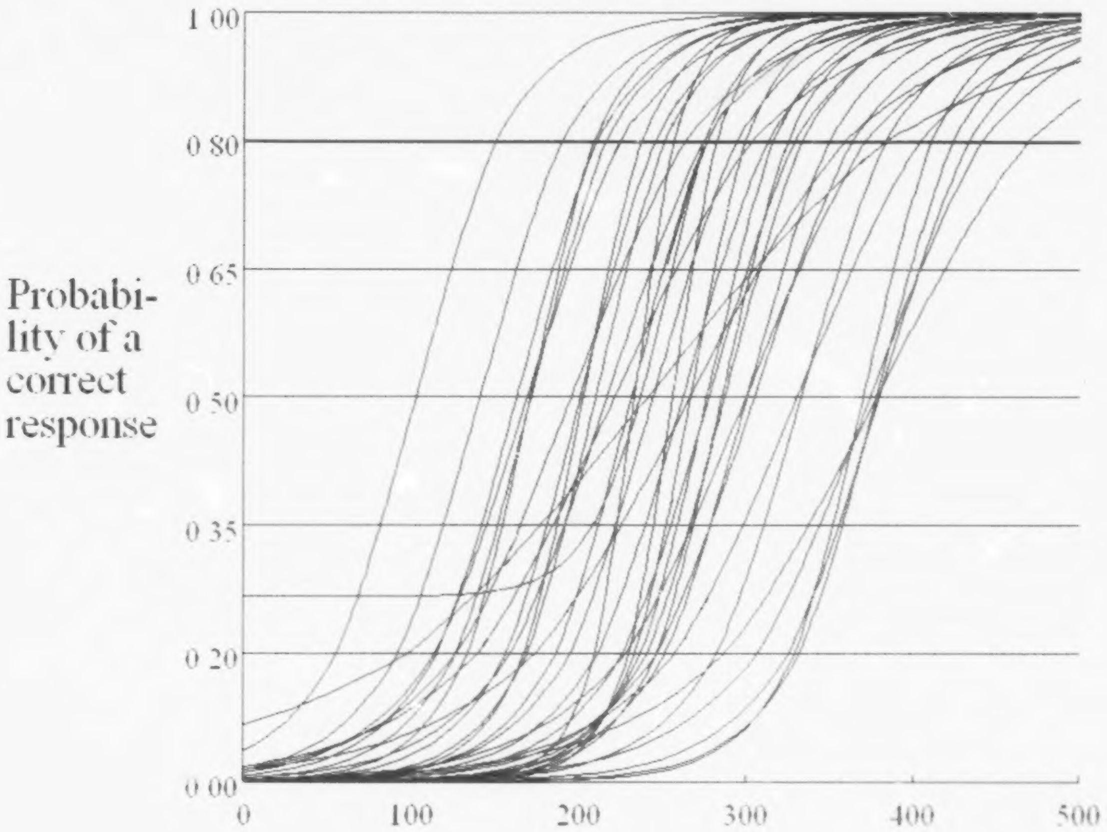


Literacy Survey. Proficiency scores below 269 indicate less than an 80 percent likelihood of success. Still, scores greater than 225 indicate an expectation for performance that is more often than not successful (i.e., greater than 50 percent). If the standard for associating items with the scale were set at 35 percent, an average task would be associated with a score of 205 or more. The score associated with any given literacy task can vary greatly, depending on the response probability adopted as the standard.

The item characteristic curves for 34 of the prose literacy tasks in the National Adult Literacy Survey are displayed in Figure 14-3. (The 7 left out would have been near the middle in difficulty and were omitted to keep the figure legible.) The ICCs are spread along the horizontal axis by empirical differences in their "difficulty." The more difficult a task, the smaller the likelihood of a correct answer at every point along the proficiency scale.

While the entire curve can be thought of as describing the difficulty of a task, simpler measures of task difficulty are often useful for summary purposes. The "difficulty parameter" in a 3-parameter IRT model indicates the point at which the item characteristic curve changes inflection and discriminates small differences in proficiency best. When it is not possible to guess the correct answer from a short list of options, this "difficulty parameter" corresponds to a 50 percent chance of success on the task. The 1992 National Adult Literacy Survey used a different measure of task difficulty—the scale score associated with reaching a response probability of 80 percent. The "RP80 difficulty" is represented graphically in Figure 14-3 by the point at which the .80 horizontal guideline intersects each ICC curve.

Figure 14–3. Item characteristic curves of 34 prose literacy tasks



In a 1-parameter IRT model, the ICCs are parallel, but the 1992 National Adult Literacy Survey used a 3-parameter IRT model to improve the fit of the model to the data. As a result, the ICCs in Figure 14-3 are not parallel curves. Taking a horizontal section of the ICCs at a different response probability, such as the .50 guideline shown in the figure, would produce a somewhat different ordering of tasks in terms of their "RP50 difficulty."

14.2 THE NEED FOR A RESPONSE PROBABILITY CONVENTION

If the slope of the item characteristic curves was vertical, it would be easy to interpret points along the prose literacy scale in term of the tasks that people at or above that point could do. The skill difference between success and failure to perform a specific task would be small. In reality, however, assessment tasks often do not provide such an absolute indication of proficiency. The improvement in performance on any given task with prose literacy is gradual, not sudden. The issue of selecting a particular value for a desired response probability arises because the slope of the item characteristic curve is less than vertical.

In the early 1980s, the National Assessment of Educational Progress (NAEP), which assesses school-aged children, developed a method of scale anchoring in order to provide descriptions of the kinds of things students know and can do at selected ranges along the NAEP proficiency scales (Beaton, 1987). Anchoring is a way to describe, in general terms, those particular assessment items at selected points along the proficiency scales for which students can succeed at least a certain percentage of the time, and for which those at the next lowest point are less successful. This procedure relies on a response probability convention. The IRT model's "difficulty parameter" roughly corresponds (except for multiple-choice questions) to a response probability convention of 50 percent. If the convention were set here, those above the boundary would be more likely to get an item right than get it wrong, while those below that boundary would be more likely to get the item wrong than right. This convention, however, was rejected and a response probability criterion of 80 percent (RP80) was chosen, in order to ensure that students above this criterion would have a sufficiently high probability of success on individual items.

The 1985 Young Adult Literacy Assessment was conducted as a part of NAEP, and included one of the 1984 NAEP reading assessment blocks. In order to anchor the literacy scales, the ETS analysts carried over the NAEP RP80 criterion for its reporting (Kirsch, Jungeblut, and others, 1986). The RP80 criterion was subsequently carried over to the 1992 National Adult Literacy Survey, in order that the adult literacy findings remained comparable with the findings from the 1985 Young Adult Literacy Assessment.

The RP80 difficulty criterion was chosen not only because NAEP had been using this criterion at the time of the 1985 survey, but also because the 1992 National Adult Literacy Survey rejected the IRT model's "difficulty parameter" as the point along the scale that should be used to differentiate success and failure. The survey analysts rejected this measure because they believed that examinees with a 50 percent chance of success (the probability associated with the "difficulty parameter") have not demonstrated sufficient mastery of that item. In order to have sufficient confidence that adults of a particular prose proficiency were able to succeed with each task, the 1992 National Adult Literacy Survey adopted NAEP's 80 percent response probability as a criterion. This convention was adopted in order to reduce the proportion of false positives[1] and to assure the accuracy of claims that examinees can meet the demands of the underlying scale at the point at which the literacy task is associated with RP80.

However, during the intervening years between 1985 and 1992, NAEP changed its response probability criterion from 80 percent to 65 percent. Eugene Johnson, the NAEP technical director, described the reasons that NAEP adopted the RP65 convention in an internal ETS memo (1994):

---

[1] A "false positive" means that an examinee has a scale score above an item's location, implying a positive prediction of success on that item, but actually fails the item. A "false negative" means that an examinee has a scale score below an item's location, implying a negative prediction of failure on that item, but actually succeeds with the item.

> While the RP percentage of 65 is arbitrary, it was selected after careful consideration of the purpose: describing students' level of performance. A larger RP percentage, such as 80, would result in higher item mapping points for all items. The result would be that smaller percentages of student would exhibit performance consistent with each exercise. For example, in the 1992 writing assessment, using a RP percentage of 65 resulted in most writing tasks having the highest score category being mapped onto the scale well above the proficiency levels exhibited by the vast majority of the assessed population of students. If an RP percentage of 80 had been used, this would likely have been true for both of the two highest score categories. In contrast, a smaller RP percentage, such as 50, would lower the mapping criteria to only a 50/50 chance that students at the scale point could provide the responses of the quality described on the map. The RP value of 65 was selected as an intermediate value to describe students' level of performance since it corresponded to a reasonably high probability of success on the questions while better matching the observed performance of the assessed population.

Johnson also pointed out in his memo that the public needs to be informed about the criterion level and to understand that the skills ascribed to students are predicated on the degree of success selected.

During 1994 and 1995, NAEP's Design and Analysis Committee reconsidered the appropriateness of NAEP's response probability convention. In 1995, NAEP adopted two related response probability conventions: 74 percent for multiple-choice questions (to correct for the possibility of answering correctly by guessing) and 65 percent for constructed response questions (where guessing is not a factor). Some support for the dual conventions was provided by Huynh (1998) in a paper originally drafted for NAEP's Design and Analysis Committee. Huynh decomposed the item information into that provided by a correct response and that provided by an incorrect response. Huynh showed that the item information provided by a correct response to a constructed-response item is maximized at the point along the scale at which two-thirds of the students get the question correct. (For multiple-choice questions, information is maximized at the point at which 74 percent get the question correct). Correspondingly, the item information provided by an incorrect response is maximized near the point along the scale at which one-third of the students get the question wrong. It should be noted, however, that maximizing the *item* information, rather than the information provided by a *correct response*, would imply an item mapping criterion closer to 50 percent.

While Huynh's analyses were influential, NAEP's dual response probability conventions (65 and 74 percent) were based, in part, on an intuitive judgment that they would provide the best picture of reading skills for students at particular points on the reading scales. The National Adult Literacy Survey staff also made an intuitive judgment, but a judgment that was more stringent. Other testing programs continue to use the IRT model's "difficulty parameter," a less stringent standard, as the criterion for locating items.

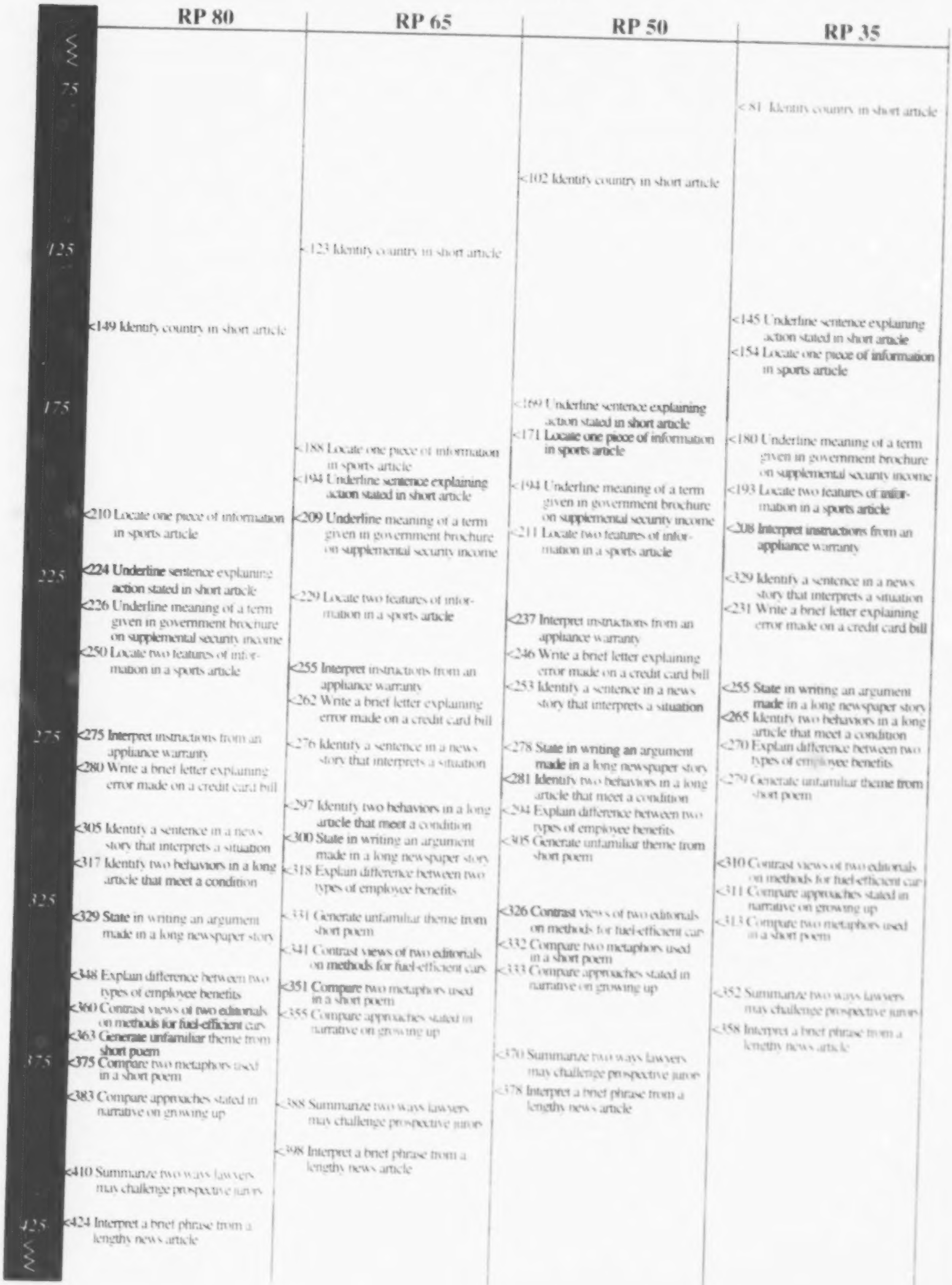
The principal uses of the response probability convention are for item mapping and scale anchoring. The 1994 NAEP reading report, for example, provided an item map that graphically displays the relative difficulty of a select group of reading assessment items. The authors (Campbell, Donahue, Reese, and Phillips, 1996) explained the meaning of their item map in the following terms:

> Each reading question was mapped onto the NAEP literacy subscale based on students' performance. The point on the subscale at which a question is positioned on the map represents the subscale score attained by students who had a 65 percent probability of successfully answering the question. Thus it can be said for each question and its corresponding subscale score—student with proficiency scores above that point on the subscale have a greater than 65 percent chance of successfully answering the question, while those below that point have a less than 65 percent chance. (The probability was set at 74 percent for multiple-choice items.)

The first report from the 1992 National Adult Literacy Survey also presented an item map to aid in the interpretation of each literacy scale (Figure 1, page 10). However, little interpretation was provided and the report did not mention the response probability convention used to map the literacy tasks. The report indicated only that this figure "describes some of the literacy tasks and indicates their scale values." An explanation like that above would also be applicable to the prose literacy scale, except that the probability used was 80 percent rather than 65 percent, and the same value was used for all items, including multiple-choice items.

In order to begin an exploration of the impact of the response probability convention on reporting data from the 1992 National Adult Literacy Survey, new item maps were developed for comparison with the item map in the initial survey report. Figure 14-4 displays item maps prepared using response probability conventions of 80 percent, 65 percent, 50 percent, and 35 percent. The left column of Figure 14-4 displays the item map for the 1992 prose literacy scale, using the 80 percent criterion as published in the initial report. Three additional columns in the figure display where the prose literacy tasks would be located had the items been mapped at 65 percent, 50 percent, or 35 percent. The literacy tasks slide up or down the proficiency scale, depending on the chosen value of the response probability criterion.

Figure 14-4. Difficulty values of selected tasks along the prose literacy scale, mapped at four response probability criteria: The 1992 National Adult Literacy Survey



BEST COPY AVAILABLE

NCES's various programs for assessing the skills of children and of adults have set or changed their response probability conventions for reasons unique to the needs of each study with no attempt to maintain a common standard. As a result, the 1992 National Adult Literacy Survey is now left with a more stringent standard for describing the skills of adults than that used to describe the skills of elementary and secondary school children. The response probability convention plays a significant role in deciding how much ability is needed to qualify as "able to do" some prose literacy task. It is not widely understood how this little-noticed convention fundamentally affects the measurement of the proportions of adults that meet the requirements of the various literacy levels. The next section describes how the response probability convention plays a role in the descriptions of levels of prose literacy used in literacy assessment surveys.

14.3 LITERACY TASKS AND LITERACY LEVELS

The prose literacy tasks in the 1992 National Adult Literacy Survey were developed in order to simulate the everyday literacy activities that people engage in when they use printed materials, and to require of adults the same literacy skills that adults normally encounter in occupational, community, and home settings. Each literacy task consisted of two parts: a selection of printed material, and a request to do something that indicated the adult could use the information contained in that material. The degree of literacy needed to successfully complete the assessment tasks is derived from three factors: 1) the format of the printed material, 2) the content of the material, and 3) the nature of the request requiring use of the material. The more difficult the literacy task, the higher the literacy skill needed to successfully complete it. Analyzing the sources of the difficulty of literacy tasks helps to understand the nature of literacy skills.

As described in Chapter 13, Kirsch, Mosenthal, and Jungeblut developed a system for measuring different aspects of the cognitive demands that literacy tasks place upon readers: type of match, plausibility of distractors, abstractness of information, and readability of the text. The 41 prose literacy tasks in the 1992 National Adult Literacy Survey were scored. The resulting distribution of scores for each of the skill and process demands were displayed previously in Tables 13-2, and 13-3.

As indicated in Tables 14-1 and 14-2, the most frequent scores on type of match were 3, 4, and 6, with only six tasks in the two easiest categories. The most frequent score on plausibility of distractors was a 2, with eight tasks having no distractors at all (score of 1). The most frequent score on abstractness of information was a 4, with six tasks in the most concrete category (score of 1). The texts used ranged widely in readability, with 4 tasks at grade levels 4 and 5, and 5 tasks at a grade level beyond high school (grade 12).

Table 14-1. Distribution of predictor scores for 41 prose literacy tasks

Score	Type of Match	Plausibility of distractors	Abstractness of information
	Number of Tasks		
1	5	8	6
2	1	18	7
3	11	3	9
4	10	8	14
5	5	4	5
6	9	0	0

Table 14-2. Distribution of readability for 41 prose literacy tasks

Grade Level	Readability of text
Number of Tasks	
4	3
5	1
6	8
7	6
8	8
9	5
10	5
13	2
15	3

If much of the variability in task difficulty can be predicted by these skill and process requirements of the cognitive tasks, then it becomes possible for these skill and process requirements to provide the building blocks for general descriptions of the kinds of skills needed to score in the selected ranges along the literacy scales that constitute the literacy levels. Describing the literacy levels in terms of general skills, rather than specific tasks, is useful because the descriptions can convey the content of the literacy scales in general, rather than how literacy skills are brought to bear on particular literacy tasks.

14.3.1 Predictive Factors and RP80 Task Difficulty

In the preceding chapter, Kirsch, Mosenthal, and Jungeblut used multiple regression to predict the difficulty of the prose literacy tasks in the 1992 National Adult Literacy Survey on the basis of the skill and process requirements of the cognitive tasks described above: type of match, plausibility of distractors, abstractness of information, and readability of the prose text. Using the prose scale scores evaluated at an 80 percent response probability convention, Kirsch and Mosenthal obtained the following estimates of the regression coefficients ($R^2 = .87$):

$$\begin{array}{lllll} \text{RP80} = 28.9\ \text{TypMatch} & +16.1\ \text{Distract} + & 8.8\ \text{Abstract} + & .2\ \text{Readability} + \text{Constant} \\ \quad (3.4)\ \text{std.err.} & (3.6)\ \text{std.err.} & (4.2)\ \text{std.err.} & (1.7)\ \text{std.err.} \end{array}$$

This equation showed that while prose task difficulty was highly predictable by these four factors, 'type of match' had a large significant impact (more than 8 times its standard error); 'plausibility of distractors' had a significant impact (more than 4 times its standard error); and 'abstractness of information' also had a significant impact (more than 2 times its standard error). Readability of the text was not an important factor in explaining task difficulty, after controlling for the other predictors.

The three authors divided the tasks into levels based on their observation of qualitative shifts in these skill or process requirements of the cognitive tasks, which they had shown were associated with increasing task difficulty. The RP80 criterion provided a method for locating the tasks on the literacy scale, while the observation of qualitative shifts in cognitive requirements provided a method for grouping the tasks into five levels based on similarities in their demands. Both grouping the tasks on the basis of their cognitive demands and locating the tasks on the scale on the basis of the RP80 criterion provided the basis for choosing the boundaries between the five literacy levels. Kirsch, Jungeblut, and Mosenthal described how they grouped the tasks on the basis of qualitative shifts in cognitive demands as follows:

> …there appears to be an ordered set of information-processing skills and strategies that may get called into play to accomplish the range of tasks represented by the three literacy domains.
>
> …As tasks moved up the scales (i.e., became more difficult), the associated [scores on the three factors] also increased. This relationship between [RP80] task difficulty and [scores on the three factors] appeared to be quite systematic. That is, toward the bottom of each literacy scale the [score on the three factors] of 1 was dominant, [scores] of 2 and 3 became more frequent as tasks move up the Prose, Document, and Quantitative Scales, and toward the higher end [scores on the three factors] of 4, 5 and higher became predominant. Although the patterns differed somewhat from scale to scale reflecting differences in the [scores on the three factors], the points on the scale at which major shifts in the processes and skills required for successful task performance were remarkably similar.

To locate the tasks on the literacy scales, Kirsch, Jungeblut, and Mosenthal relied on the 80 percent response probability criterion. The three authors observed several transition points (occurring at roughly 50 point intervals and beginning with 225 on each scale) at which the aggregate task demands shifted upwards. They decided that using equal-interval cutpoints that were the same on each scale (225, 275, 325, and 375) constituted a discretionary degree of rounding that was consistent with the accuracy of predicting the RP80 difficulty of the tasks. They divided each scale into five levels that reflect both the 80 percent response probability criterion and the progression of information-processing skills and strategies: Level 1 (0-225), Level 2 (226-275), Level 3 (276-325), Level 4 (326-375), and Level 5 (376 to 500).

Once the cut points between the levels were decided, Kirsch and Mosenthal wrote general descriptions of the kinds of demands placed on readers by tasks in each of their five levels (Table 13-13). The descriptions attempted to capture the various combinations of the three important predictors of difficulty among typical tasks at each of the five literacy levels.

There are some minor problems with these descriptions. No description of the 'abstractness of information' variable was included in the most well-known version of the level descriptions (the middle column), even though it had a significant impact on task difficulty. In addition, the description of Level 1 includes an unwarranted term—"relatively short text"—that describes the readability of the prose stimulus, a factor that their regression analysis showed was not essential to item difficulty when the other factors were included. Nevertheless, the descriptions of literacy levels have served the purpose of providing a general summary of the skills required to succeed (at the 80 percent level) on the literacy tasks included in the 1992 National Adult Literacy Survey.

14.3.2 Predictive Factors, Task Difficulty, and the Response Probability Convention

Kirsch, Mosenthal, and Jungeblut conducted all their analyses using task difficulty as measured at the RP80 response probability convention, and then used their results to justify grouping the literacy tasks into levels for summary desrcriptive purposes. In order to isolate the impact of the response probability convention on reporting results from the 1992 National Adult Literacy Survey, it is necessary to perform similar analyses, using other response probability conventions as a criterion for task difficulty. The new analyses could then indicate whether the prior descriptions of literacy levels would remain appropriate when other response probability conventions are used.

Table 14-3 shows estimates of the corresonding multiple regression coefficients using task difficulty measured at the original 80 percent and fourteen alternative response probability conventions, from 20 to 90 percent. For these regressions, the number of cases was expanded to 71 by including 30 additional prose literacy tasks that had been used in the 1991 study of the literacy of job-seekers (Kirsch, Jungeblut, and Campbell, 1992).

The coefficients in Table 14-3 display several patterns that could not be seen in a single regression with RP80 as an outcome. The coefficient of explained variance (R^2) increases as the response probability falls from RP90 to RP60, then decreases with lower response probabilities. At high response probability levels, the importance of the 'plausibility of distractors' factor is greatest and readability is not a significant factor. However, the importance of these two factors reverses at low response probability levels. At RP35 and below, the coefficient of 'plausibility of distractors' goes below twice its standard error and becomes insignificant. At RP55 and below, the coefficient of readability becomes a significant factor in explaining task difficulty. The increase in variance explained and the fact that readability becomes a significant factor at lower RP criteria may suggest that there is potentially useful information to be captured in describing literacy levels should a lower criterion be used for reporting.

The boundaries between the prose literacy levels were based on two factors: a clustering of prose tasks with similar cognitive demands and the locations of the literacy tasks on the scale based on the RP80 criterion. In order to isolate the impact of the response probability convention on reporting results from the 1992 National Adult Literacy Survey, it is necessary to find cut points between the levels that will group the literacy tasks together on the same basis, with the same qualitative shifts in the cognitive demands of the tasks, as measured by the four task variables that Kirsch, Mosenthal, and Jungeblut reported. Since the item characteristic curves of the tasks are not parallel, some tasks change their ordering slighlty from one response probability convention to another. It is not possible to divide exactly the same tasks into the exact same five levels, just by changing the response probability criterion. It is possible, however, to use the regression equations to derive alternative cut points between levels that would approximate the same grouping of prose literacy tasks by skill and process demands. The next section describes the process used to derive alternative cutpoints that would ensure, to the maximum extent possible, that the alternative cutpoints reflect only the influence of changing the response probability convention and not other, more substantive considerations.

Table 14-3. Multiple regression coefficients, standard errors, and R-squares for regression equations predicting task difficulty measured at selected response probability criteria for 71 prose literacy task from the 1992 National Adult Literacy Survey and the 1991 study of the literacy of job-seekers

Response probability criterion	Inter-cept	Type of match	(Std. Error)	Plausibility of distractors	(Std. Error)	Abstractness of information	(Std. error)	Readability of prose text	(Std. error)	R^2
0.90	157.2	21.8	(3.4)	21.2	(3.7)	12.4	(4.1)	0.2	(1.9)	0.765
0.85	144.2	22.1	(3.1)	18.6	(3.3)	11.9	(3.7)	0.8	(1.7)	0.794
0.80	134.3	22.3	(2.8)	16.5	(3.0)	11.5	(3.4)	1.2	(1.6)	0.813
0.75	126.2	22.5	(2.6)	14.9	(2.8)	11.1	(3.2)	1.6	(1.5)	0.826
0.70	119.1	22.6	(2.5)	13.4	(2.7)	10.8	(3.0)	1.9	(1.4)	0.835
0.65	112.7	22.8	(2.4)	12.1	(2.6)	10.6	(2.9)	2.2	(1.4)	0.839
0.60	106.7	22.9	(2.4)	10.8	(2.6)	10.3	(2.9)	2.5	(1.4)	0.841
0.55	100.9	23.1	(2.4)	9.6	(2.6)	10.0	(2.8)	2.8	(1.3)	0.840
0.50	95.2	23.2	(2.4)	8.5	(2.6)	9.8	(2.8)	3.0	(1.3)	0.837
0.45	89.6	23.4	(2.4)	7.3	(2.6)	9.5	(2.9)	3.3	(1.4)	0.830
0.40	83.7	23.6	(2.5)	6.1	(2.6)	9.2	(2.9)	3.6	(1.4)	0.821
0.35	77.6	23.8	(2.5)	4.8	(2.7)	8.9	(3.0)	3.8	(1.4)	0.809
0.30	70.5	24.1	(2.7)	3.4	(2.9)	8.5	(3.2)	4.2	(1.5)	0.792
0.25	65.4	24.3	(2.8)	1.8	(3.0)	8.0	(3.4)	4.3	(1.6)	0.767
0.20	64.5	25.3	(3.1)	-0.5	(3.4)	6.8	(3.8)	3.9	(1.8)	0.716

14.3.3 Alternative Cut Points between Literacy Levels

The goal of deriving alternative cut points between levels first requires developing a general method that reproduces the existing RP80 cut points, then extending the method to other RP values. Since the cut points are derived by grouping similar tasks together, developing sets of hypothetical tasks with difficulty-related characteristics similar to those in the existing literacy levels appeared to be a workable approach. Table 14-4 displays the values of a selected group of hypothetical tasks—about half a dozen for each level. These are "literacy tasks" only in the sense that if real literacy tasks were developed that had the selected values on the variables measuring cognitive demands and that showed the same relationship to task difficulty as the 71 existing tasks used in the regression analyses, the sets of tasks in each level would display the following properties:

- When used to predict the average RP80 task difficulty within each level for the set of hypothetical tasks (using the RP80 equation in Table 14-3 and averaging the predicted values within each level), the predicted averages would come to exactly the midpoints of each of the five existing literacy levels: 200, 250, 300, 350, and 400.
- When the scale distance between pairs of adjacent midpoints are divided in half, the results become the existing RP80 cutpoints between levels: 225, 275, 325, and 375.

- When the hypothetical items are used to predict average task difficulty within each level for alternative response probability criteria, the scale distances between pairs of adjacent midpoints provide alternative cut points between levels, based on the other response probability criteria.

The five sets of hypothetical literacy tasks, along with their scores on the values of the variables affecting task difficulty for each of the hypothetical tasks are shown below in Table 14-5. The values of the task variables were multiplied by the various sets of multiple regression coefficients (shown in Table 14-3) associated with each response probability criterion and averaged within levels to produce new sets of midpoints. The points on the scale half-way between the sets of midpoints provided alternative cut points between levels, and are shown in Table 14-4.

Table 14-4. Alternative cut points between prose literacy levels, by response probability criteria: The 1992 National Adult Literacy Survey

Response probability criterion	Between prose literacy levels			
	1 and 2	2 and 3	3 and 4	4 and 5
RP90	250	302	355	408
RP85	236	287	338	389
RP80	225	275	325	375
RP75	216	266	315	364
RP70	209	258	305	354
RP65	202	250	297	345
RP60	195	243	289	336
RP55	189	237	282	328
RP50	183	230	274	320
RP45	177	223	267	312
RP40	171	217	260	303
RP35	164	210	252	295
RP30	157	202	243	285
RP25	150	194	234	275
RP20	142	185	223	261

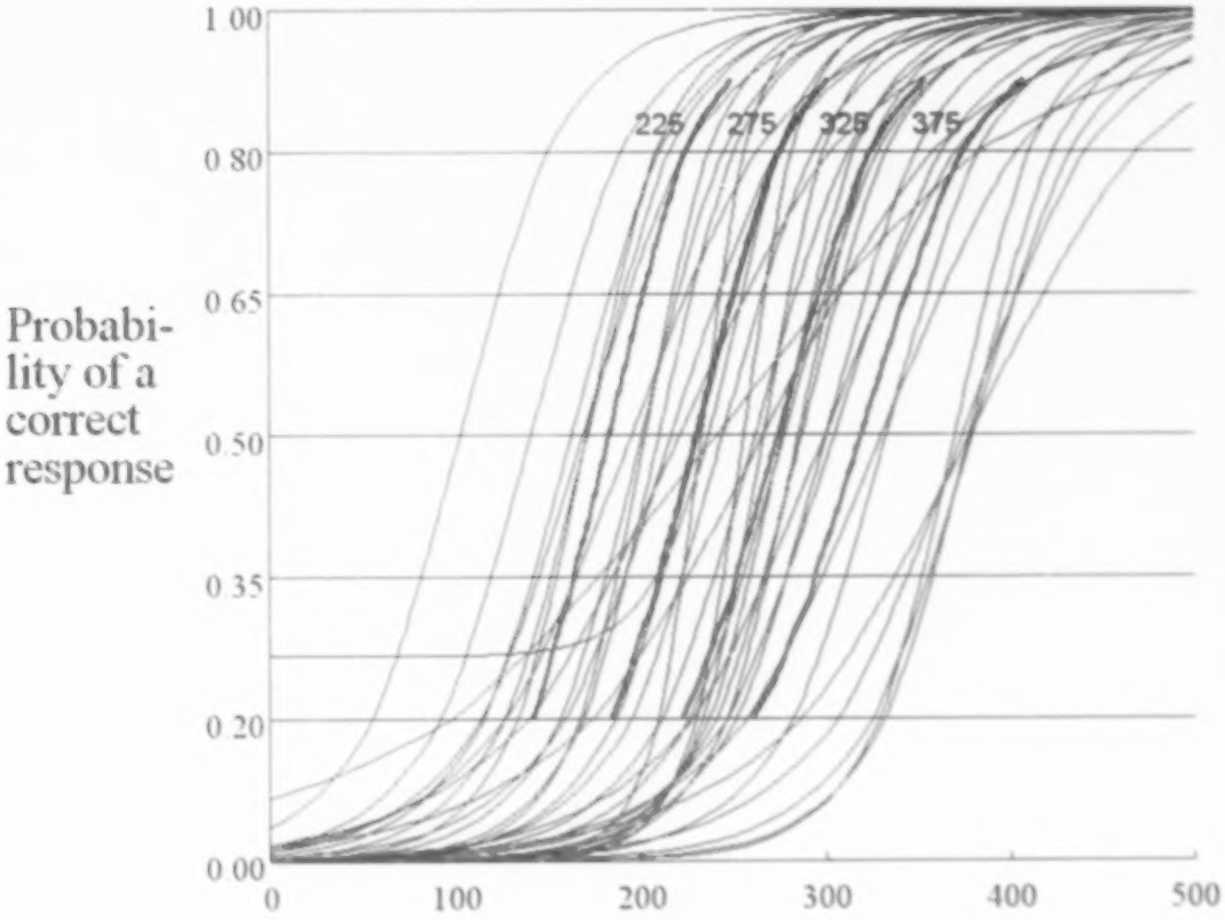
Table 14-5. Scores of hypothetical prose literacy tasks on four factors that predict difficulty

Prose literacy levels	Type of match	Plausibility of distractors	Abstractness of information	Readability in grade levels
Level 1	1	1	1	4
	1	1	1	5
	1	1	1	6
	1	1	2	6
	1	1	2	5
	1	1	3	4
	1	2	1	5
Level 2	2	1	2	6
	2	2	1	7
	2	1	3	6
	2	2	2	7
	3	2	2	6
	3	1	3	7
	3	2	2	8
	2	3	3	7
Level 3	3	3	2	7
	3	3	2	8
	4	2	3	7
	4	2	3	8
	4	3	3	8
	4	3	3	9
Level 4	4	4	3	7
	4	4	4	8
	5	3	4	9
	5	4	2	9
	5	4	3	8
	5	4	3	9
	6	2	4	10
Level 5	4	5	5	13
	6	4	3	10
	5	5	4	8
	6	3	5	10
	6	4	5	9
	7	4	4	10
	6	5	5	13
	7	5	4	10

Table 14-4 shows how the alternative cut points change as a function of the response probability criterion. Around 50 percent, each increase of five points in the response probability criterion moves the cutpoints between literacy levels about 5 to 8 points higher on the prose literacy scale. Around 80 percent, each increase of five points in the response probability criterion moves the cutpoints between literacy levels about 12 to 14 points higher.

It is instructive to examine the alternative cut points between levels graphically, as well as numerically. Figure 14-4 below plots both the item characteristic curves previously shown in Figure 14-3 and the alternative cut points between the literacy levels as a thick line (with a spline interpolation between the points). Also displayed in this graph are numbers showing the four existing cutpoints along the RP80 horizontal gridline (225, 275, 325, and 375). Visual inspection of this graph leads to the conclusion that

Figure 14-4. Item characteristic curves of 34 prose literacy tasks and literacy level cutpoints between the tasks: The 1992 National Adult Literacy Survey



the majority of prose literacy tasks stay within the same level, regardless of the response probability value used. As a result, the Kirsch-Mosenthal-Jungeblut descriptions of what tasks in Levels 1 through 5 require of adults do not differ much by the response probability convention adopted, at least for criteria above RP60. If a criterion below RP60 were used, the descriptions would have to be revised to add elements relating to the readability of the text. If a criterion below RP40 were used, the descriptions would have to be revised to remove elements relating to the plausibility of distractors.

The conclusions that on the one hand, the Kirsch-Mosenthal-Jungeblut descriptions of what tasks in Levels 1 through 5 require of adults do not vary as a function of the response probability criterion, and that on the other hand, the cut points between the levels change drastically, appear to be in conflict. The process of setting alternative cut points has isolated the impact of the response probability convention on setting boundaries between the literacy levels. The current descriptions of the literacy levels, which have carried over into later reports from the International Adult Literacy Surveys, appear to lack a description of the response probability criterion and an understanding of its significance.

The variation in the cut points between levels as a function of the response probablity convention are quite large. To put the sensitivity to this factor in perspective, consider that if the 1992 National Adult Literacy Survey were to adopt the NAEP convention for mapping items onto latent scales (65 percent), the cut points between levels would have to move down by 23 to 28 points, which is about half the 50-point width of a literacy level. The standard error, a measure that incorporates variability due to both statistical sampling and measurement errors, for average prose literacy for the U.S. population as a whole was 0.7 points on the same literacy scale. The variation in the cut points as a function of the response probability convention plays a significant role in estimating the distribution of the adult population over the five literacy levels.

14.4 DISTRIBUTION OF ADULTS AMONG ALTERNATIVE LITERACY LEVELS

The initial report of the 1992 National Adult Literacy Survey found that 21 percent of the 191 million adults in this country demonstrated skills in the lowest, and another 26 percent demonstrated skills in the second lowest of five prose literacy levels, using the RP80 response probability The initial report explained that most adults in Level 1 were able to perform simple, routine tasks involving brief and uncomplicated texts and documents. For example, they were able to identify a piece of specific information in a brief news article. Others in Level 1 attempted to perform the literacy tasks and were unsuccessful. There were individuals who had such limited skills that they were able to complete only part of the survey (Kirsch, Jungeblut, Jenkins, and Kolstad, 1993). Of those who scored in Level 1, 21 percent of adults did not perform a single prose literacy task correctly (Table A.5P in Haigler, Harlow, O'Connor, and Campbell, 1994).

Once the alternative cut points between the prose literacy levels were determined, it was possible to estimate the proportion of U.S. adults who performed in each level under alternative choice of response probability conventions. The results are shown in Table 14-6 below. Each row in Table 14-6 presents the population distribution of adults across the five prose literacy levels. The rows differ only in the response probability convention used to set the cut points between the levels. For response probability conventions above 60 percent, the same general descriptions of literacy levels can be used. The only difference is the proportion of time that adults have to be successful with equivalent tasks in order to be counted as "able to do" such tasks.

Table 14-6. Percentages of U.S. adults within each level of prose literacy, defined by alternative response probability values: 1992

Response probability	Level 1		Level 2		Level 3		Level 4		Level 5	
	Prcnt	(st. err.)	Prcnt	(st. err.)	Prcnt	(st. err.)	Prcnt	(st. err.)	Prcnt	(st. err.)
90	32	(0.5)	33	(0.7)	27	(0.4)	8	(0.3)	1	(0.1)
85	25	(0.4)	30	(0.5)	31	(0.5)	13	(0.4)	2	(0.2)
80	20	(0.4)	27	(0.6)	32	(0.7)	18	(0.4)	3	(0.2)
75	17	(0.4)	24	(0.6)	32	(0.8)	21	(0.4)	6	(0.3)
70	15	(0.4)	21	(0.5)	31	(0.6)	24	(0.4)	9	(0.3)
65	13	(0.4)	19	(0.5)	30	(0.6)	26	(0.4)	12	(0.4)
60	12	(0.3)	17	(0.5)	28	(0.7)	28	(0.6)	16	(0.4)
55	10	(0.3)	15	(0.4)	26	(0.5)	29	(0.7)	20	(0.5)
50	9	(0.3)	13	(0.3)	24	(0.6)	30	(0.7)	24	(0.5)
45	9	(0.2)	11	(0.3)	22	(0.6)	29	(0.6)	29	(0.5)
40	8	(0.3)	10	(0.3)	20	(0.5)	29	(0.6)	34	(0.5)
35	7	(0.3)	8	(0.3)	18	(0.5)	27	(0.6)	40	(0.5)
30	6	(0.3)	7	(0.3)	15	(0.5)	25	(0.6)	46	(0.6)
25	6	(0.2)	6	(0.2)	13	(0.4)	23	(0.5)	53	(0.6)
20	5	(0.2)	5	(0.2)	10	(0.3)	19	(0.5)	62	(0.6)

Source: U.S. Department of Education, National Center for Education Statistics, the National Adult Literacy Survey, 1992

If the 1992 National Adult Literacy Survey had reported the same results using the somewhat lower RP65 response probability convention that is currently used in reporting the educational achievement of our nation's children in the National Assessment of Educational Progress, the report would indicated that the two lowest levels included only 13 percent and 19 percent of the 191 million adults in this country. If the 1992 National Adult Literacy Survey had reported the same results using 50 percent as a criterion, the combined proportion in Levels 1 and 2 would be reduced to 22 percent. As the criterion response probability is relaxed in Table 14-6, larger proportions of adults appear to be able to perform at higher levels of prose literacy. The response probability convention makes the most difference at the upper and lower ends of the scale (Levels 1 and 5). As the response probability convention drops from 80 to 50 percent, the proportion of the population in prose literacy Level 5 increases from 3 to 24 percent, a substantively and statistically significant change.

A factor that has such a large impact on the results of the survey deserves a thorough understanding of the issues and debate over the standard to be adopted. People concerned with measuring literacy accurately need to understand what the response probability convention is and why it matters to reporting the results.

14.5 SELECTING AN APPROPRIATE RESPONSE PROBABILITY CONVENTION

These changes in the distribution across the five literacy levels do not mean that people have more or less literacy skills than previously reported. The underlying skills of the population have not changed. What has changed is the dividing line between those who are said to be "able to do" the prose literacy tasks and those who are not.

The RP80 response probability criterion was adopted for the 1992 survey for several reasons. A sister project at Educational Testing Service, the National Assessment of Educational Progress (NAEP) had been using this criterion at the time of the 1985 Young Adult Literacy Assessment, and the results of 1992 survey were intended to be as comparable to the 1985 assessment as possible. Nevertheless, the traditional argument for a high response probability convention is not satisfactory. The issue is unrelated to the way the data were collected or the way the responses were scored, because the response probability convention is used only during the analysis and reporting stage of the survey. Findings from earlier studies could be reanalyzed using a different convention, yet retain comparability.

The National Adult Literacy Survey analysts rejected the IRT model's "difficulty parameter" as the point along the scale that should be used to differentiate success and failure because they believed that examinees with a 50 percent chance of success (the marginal probability associated with the "difficulty parameter") have not demonstrated sufficient mastery of that item. The 80 percent criterion gave the survey analysts sufficient confidence that adults at given points on the prose literacy scale were consistently able to succeed with each task. This convention was adopted in order to reduce the proportion of false positives and to assure the accuracy of claims that examinees can meet the demands of the underlying scale at the point at which the literacy task is associated with RP80. This value of the response probability criterion ensured that when survey reports indicate that adults are able to do a particular task or a generic group of tasks, the adults have mastered that task.

Unfortunately, the substantive argument for a high response probability convention is not very fine-grained. Demonstrating task mastery requires only a response probability convention well above 50 percent, not a specific value. Other choices such as 70, 75, or 85 percent could satisfy the logic of the argument, yet would produce substantially different outcomes in terms of both the cut points between and population distribution over the five levels. In addition, the cost of a high response probability convention is an increase in the proportion of false negatives. Those that the scale predicts are unable to succeed with literacy tasks have much more than a negligible probability of actually being able to succeed. Better justification is needed than has heretofore been forthcoming for retaining the 80 percent criterion in future work.

The substantive logic of the mastery argument also contains within it the seeds of an alternative perspective. Many users of adult literacy survey findings want to focus on the *lack* of facility with printed and written material and to report on what adults *can't* do. To be sure that adults are *unable* to perform the literacy tasks, survey analysts might need to use a correspondingly low response probability criterion, such as 20 percent. Such a criterion would assure that adults who do not reach these levels of consistency in their responses have a very low chance of success with the tasks. Data users can be very sure that such a group of adults does not have the skills in question. The argument for the 80 percent convention was that a high criterion is needed to ensure mastery. A similar argument could be made that a 20 percent convention is needed to ensure task failure. When the purpose of reporting is to discuss what students or adults "can't do," there may be some value in reporting achievement according to a low response probability convention. An alternative way to estimate the number of adults who did not have the skills to perform any of the tasks in prose literacy Level 1 would be to compute the proportion of adults who failed to answer correctly a single prose literacy task in the assessment, a number that turned out to be 8.2 million, or 4 percent of the adult population (See Table A.5P in Haigler, Harlow, O'Connor, and Campbell, 1994). Table 14-6 shows that a similar proportion, 5 percent of the adult population, falls in Level 1 when the response probability convention drops to 20 percent.

A large proportion of the population occupies a middle ground between those who are consistently successful and those who are consistently unsuccessful with the assessment's literacy tasks. Those who are as likely to get a question right as to get it wrong have not mastered certain skills, but they are not unskilled, either. Consider the population at the boundary between Level 1 and Level 2. Table 14-4 showed that a score of 225 was the minimum needed to ensure at least a 80 percent chance of success with the tasks in Level 1, while the minimum needed to ensure at least a 20 percent percent chance of success was 142. Table 14-6 shows 20 percent of adults score below 225, and 5 percent score below 142. The difference between the two, 15 percent of adults, represents those for whom it remains uncertain either that they will succeed or that they will fail with the tasks in Level 1.

Using item response theory as a statistical model provides an argument for mapping items at the "difficulty parameter," which is equivalent to 50 percent for nearly all the tasks in the 1992 National Adult Literacy Survey. The main purpose of creating a scale rather than reporting success on particular items is to focus on the common skills that all the items together require. The more the response probability criterion deviates from 50 percent, the more weight is given to peripheral factors unrelated to the core skill requirements that the items share in common (Kolstad, Cohen, Baldi, Chan, DeFur, and Angeles, 1998).

Because the tasks used in the assessment are not perfect measures of literacy, there may be a substantial proportion of adults in the middle range, who have neither fully mastered the literacy tasks nor are completely unable to succeed with them. The value of a 50 percent response probability convention is that it strikes an even balance between false positives and false negatives. With this criterion, one can have the same degree of confidence in statements about what adults can't do as in statements about what adults can do. Mapping items in this way both provides a balance between false positive and false negative claims about adult performance and minimizes the total misclassification error.

14.6 CONCLUSIONS

It should be reiterated that while varying the response probability convention would distribute the population differently across the five levels of literacy, adults still have exactly the same literacy skills as previously reported. The true literacy proficiencies of the population are unaffected by variations in the way analysts interpret the data. What changes is the dividing line between those classified as "able to do" the literacy tasks and those not classified as "able to do" the literacy tasks. The substantive argument for the highest possible response probability convention was that maximum practical mastery is needed to describe readers accurately as "able to do" the literacy tasks.

Several alternatives are possible. One would be to use a fixed "mastery" increment (set using professional judgment) to the the IRT model's "difficulty parameter" to satisfy the mastery criterion, rather than a fixed, high response probability convention. The increment would ensure that examinees have a sufficient likelihood of success, would put more weight on items that disciminate literacy better and less weight on items that discriminate more poorly, and would highlight the judgmental basis of the item mappings. Nevertheless, an imbalance between false positives and false negatives and inconsistencies with the statistical model would remain. A second approach would be to use the IRT model's "difficulty parameter" for mapping, (roughly equivalent to a 50 percent criterion). This would provide a better balance between false positives and false negatives, more consistency with other aspects of the IRT statistical machinery, and equivalent weight on all items, regardless of their discriminating power.

Perhaps what is needed is a fresh approach to interpreting the findings, one that does not rely on item mapping for interpreting the findings. Some have suggested interpreting scale scores in terms of the expected percentage correct for a standard set of released items (a "market basket" approach), as a way of escaping the arbitrariness of choosing a particular value for the mastery increment or the response probability convention. Another alternative, one adopted by the National Assessment of Educational Progress, is to use judgment-based standards for reporting. The National Assessement Governing Board's achievement level setting process, though it has been criticized, does not involve item mapping and would provide a judgmental standard for how much literacy is sufficient, something that the public seems to want.

Chapter 15

WORKING WITH SPSS AND SAS

Patricia O'Reilly, Educational Testing Service

This chapter discusses the use of the statistical software SPSS and SAS for analyzing the 1992 National Adult Literacy Survey data. Included are procedures for creating SPSS and SAS system files, and a description of how to use the jackknife procedure with SPSS and SAS to estimate standard errors. The SAS code works on SAS version 6.0 and higher. The SPSS code works on SPSS version 8.0.

15.1 The Electronic Code Book and SPSS and SAS Control Statements

The data files derived from the 1992 National Adult Literacy Survey are available from the National Center for Education Statistics and can be accessed using an electronic code book produced by NCES staff. The electronic code book is available in a Windows version (Windows '95 and '98). One of the options available in the electronic code book is to produce SPSS and SAS control statements to facilitate the creation of SPSS and SAS system files. These control statements include statements for variable definitions, variable labels, missing value codes, value labels, and an optional section for creating and storing scored variables. The electronic code book is arranged in sections with variables for the household sample first, variables for the prison sample next, and the variables for the non-incentive sample last. The data for each of these samples is stored in a separate file. The electronic code book also contains unweighted descriptive statistics of the reporting variables for the three related data files (see Appendix P for a more comprehensive guide for use of the electronic code book).

The features produced by the code book for the control statements include the following items.

VARIABLE DEFINITION	The field names are listed in the order in which they appear on the file. The electronic code book will produce INPUT or DATA LIST statements with column positions and input formats.
VARIABLE LABELS	A 40-character text description for each field.
MISSING VALUES	Missing values in the data are sometimes represented by blank fields. However, others have special codes for "I Don't Know," omitted, not reached, or multiple responses.
VALUE LABELS	All numeric fields with discrete (or categorical) values are provided with 20-character text descriptors for each value within the variable's range.
SCORING	Correct responses are identified in the electronic code book with an asterisk. Correct response values are usually scored as 1; all other legitimate values, including no response and "I don't know" are usually scored as 0. Values that are missing for various reasons are identified separately, so that the user can substitute alternate values.

15.2 Creating SPSS System Files

The standard structure of an SPSS program to process the 1992 National Adult Literacy Survey data should be generated according to the structure in Table 15-1. This is the format of the SPSS file generated by the electronic code book.

Table 15-1. SPSS control statement synopsis

```
SET COMPRESSION=ON.
TITLE
     Label for sysout of file generation run

FILE LABEL
     label to be stored with file

DATA LIST FILE=RAWDATA
     variable names, locations, and formats

VARIABLE LABELS
     40-character label for each variable

MISSING VALUES
     list of variables to have user-missing values assigned

VALUE LABELS
     variable names, values, and value labels

TEMPORARY.

RECODE
     oldvar (SYSMIS=SYSMIS) (mrval=9) (keyval [s]=1)
     (nrval=0) (omval=0) (idkval=0) (ELSE=0)
```

Table 15-1. SPSS control statement synopsis — Continued

```
MISSING VALUES
      for recodes of multiple responses


VALUE LABELS
      1=Correct 0=Incorrect


FREQUENCIES
      reporting variables


SAVE OUTFILE=SYSFILE/COMPRESSED.

DISPLAY LABELS.
```

The missing value transformations are followed by a series of RECODE scoring statements to create scored variables from cognitive item variables (see Section 15.1). The TEMPORARY command instructs SPSS to perform the subsequent RECODE statement on a temporary basis. Once the next encountered procedure (FREQUENCIES) has run, the variables created in the RECODE statement are deleted. Thus, the scoring variables will NOT be saved on the system file.

The control statements were coded according to the command and procedure descriptions in the *SPSS Reference Guide* (SPSS, Inc., 1990).

15.3 Creating SAS System Files

The standard structure of a SAS program to process the 1992 National Adult Literacy Survey data should be generated according to the structure in Table 15-2. This is the format of the SAS file generated by the electronic code book.

Table 15-2. SAS control statement synopsis

```
TITLE
DATA SYSFILE.xxx;
INFILE RAWDATA;
INPUT
      variable names, positions, and formats;

LABEL
      40-character variable labels;
```

Table 15-2. SAS control statement synopsis — Continued

```
LENGTH DEFAULT=2
     Other variables with appropriate lengths;

ARRAY DKn (I)   list of cognitive items with "I Don't Know"
* DO OVER DKn;  codes to be recoded for
                missing
* IF DKn=7 THEN KDn=.;
* END;

ARRAY OMn (I)   list of variables with omit
                codes to be
* DO OVER Omn;  recorded for missing
* IF OMn=8 THEN OMn=.;
* END;



ARRAY NRn (I)   list of variables with not
                reached codes
* DO OVER NRn;  to be recoded for missing
* IF NRn=9 THEN NRn=.;
* END;

ARRAY MRn (I)   list of variables with
                multiple response
* DO OVER MRn;  codes to be recoded for
                missing
* IF MRn=9 THEN MRn=.;
* END;


%MACRO RECODE;
                 SAS macro to perform scoring for each variable
%MEND RECODE;
%MACRO SCORE;
%RECODE          (oldvar, newvar, idkval, omval, nrval, mrval,
key1val[key2val])
     .
     .
%MEND;
*%SCORE          delete asterisk to save scored variables **
RUN;

PROC FORMAT LIBRARY=SASLIB;
 VALUE
      formats for the reporting variables;
PROC FREQ;
TABLES
      reporting variables;
PROC CONTENTS NOSOURCE POSITION;
RUN;
```

In SAS, the SAS Macro Language facility can be used to reduce the number of source statements generated and provide consistent performance of repetitive functions. However, the MACRO facility is not the default option; the user must ensure that the MACRO option is properly activitated when using control statements that invoke MACROs.

The DO OVER through END statements following each ARRAY statement can be used to set up the conversion of the "I Don't Know" omit, not reached, and multiple response codes to SAS's numeric missing value. However, once this conversion is executed and saved on the system file, these recoded values will be indistinguishable from actual missing values on the original data file.

The missing value transformations are followed by a series of SAS macro definitions for scoring the cognitive items. The RECODE macro is used by the SCORE macro to transform the responses to each item into score values. The RECODE macro may be edited by the user to consistently transform the special codes for each item into other values.

At the end of the control statement synopsis, the SCORE macro is commented out. To save the scored variables on the system file, the user should uncomment the % SCORE statement.

SAS saves format specifications in a user-specified library designated to the system as SASLIB. Included with the electronic code book for each sample is a list of all discrete variables and the format values to be used in any SAS analysis.

The control statements were coded according to the command and procedure descriptions in *SAS Language: Reference, Version 6, First Edition* (SAS Institute, Inc., 1990).

15.4 Computing the Estimated Variance of a Statistic Using Jackknife Methods in SPSS or SAS

This section presents examples, in the form of SPSS and SAS program code, of two multiweight methods for computing the estimated variance of a statistic, such as a percentage or a mean. The reader may wish to refer to chapter 9 for more information about the jackknife procedure. The first multiweight methods (Tables 15-3 and 15-4) may be used for any variable except the plausible values for the Prose, Document, or Quantitative scales. For example, this method is used to estimate the standard error of percentages. The second multiweight method, which should be implemented to calculate the standard error of proficiency means (see Tables 15-5 and 15-6), employs a correction for the variance due to estimation of the plausible values. This method should be used to estimate the standard error of proficiency means.

Table 15-3. Standard error computation: Multiweight method using SPSS

```
GET FILE=SYSFILE/     (System file for sample)
      KEEP=BG12401, WEIGHT, REPL01 TO REPL60, X
VECTOR WT=REPL01 TO REPL60.
VECTOR WX(60).
SELECT IF (NOT SYSMIS(X)).
COMPUTE WTX=WEIGHT*X.
LOOP #I=1 TO 60.
  COMPUTE WX(#I)=WT(#I)*X
END LOOP.
```

Table 15-3. Standard error computation: Multiweight method using SPSS — Continued

```
AGGREGATE OUTFILE=*/BREAK=BG12401/UWN=N(WEIGHT)/
      SWT,SW1 TO SW60=SUM(WEIGHT, REPL01 TO REPL60)/
      SWX,SX1 TO SX60=SUM(WTX, WX1 TO WX60)/.
VECTOR SW=SW1 TO SW60.
VECTOR SX=SX1 TO SX60.
COMPUTE XBAR=SWX/SWT.
COMPUTE XVAR=0.
LOOP #I=1 TO 60.
  COMPUTE#DIFF=SX(#I)/SW(#I) - XBAR.
  COMPUTE SVAR=SVAR + #DIFF * #DIFF.
END LOOP.
COMPUTE  XSE=SQRT(XVAR).
PRINT FORMATS XVAR,XSE(F8.4).
LIST VARIABLES=BG12401,UWN,SWT,XBAR,XVAR,XSE.
```

Table 15-4. Standard error computation: Multiweight method using SAS

```
DATA A;
      SET SYSFILE.MAI;
      ARRAY WT REPL01-REPL60;
      ARRAY WX WX1-WX60;
      IF (X NE .);
      WTX=WEIGHT*X;
      DO OVER WT;
            WX=WT*X;
      END;
PROC SUMMARY;
      CLASS BG12401;
      VAR WEIGHT REPL01-REPL60 WTX WX1-WX60;
      OUTPUT OUT=B N(WEIGHT)=UWN
            SUM(WEIGHT WTX REPL01-REPL60 WX1-WX60)=
              SWT SWX SW1-SW60 SX1-SX60;
DATA C;
      SET B;
      ARRAY SW SW1-SW60;
      ARRAY SX SX1-SX60;
      XBAR=SWX/SWT;
      XVAR=0;
      DO OVER SW;
            DIFF=(SX/SW)-XBAR;
            XVAR=XVAR+DIFF*DIFF;
      END;
      XSE=SQRT(XVAR);
PROC PRINT;
      VAR BG12401 UWN SWT XBAR XVAR XSE;
```

Table 15-5. Standard error computation: Multiweight method using SPSS with correction for imputation

```
GET FILE=SYSFILE/     (System file for sample)
      KEEP=BG12401, WEIGHT, REPL01 TO REPL60, NRPP11 TO NRPP15.
VECTOR VALUE=NRPP11 TO NRPP15.
VECTOR WT=REPL01 TO REPL60.
VECTOR WX(60).
VECTOR WS(5).
SELECT IF (NOT SYSMIS(NRPP11)).
COMPUTE WTX=WEIGHT*NRP11.
LOOP #I=1 TO 60.
  COMPUTE WX(#I)=WT(#I)*NRPP11.
END LOOP.
LOOP #I=1 TO 5.
      COMPUTE WS(#I)=VALUE(#I)*WEIGHT.
END LOOP.
AGGREGATE OUTFILE=*/BREAK=BG12401/UWN=N(WEIGHT)/
      SWT,SW1 TO SW60=SUM(WEIGHT, REPL01 TO REPL60)/
      SWX,SX1 TO SX60=SUM(WTX, WX1 TO WX60)/
      SS1 TO SS5 = SUM(WS1 TO WS5)/.
VECTOR SW=SW1 TO SW60.
VECTOR SX=SX1 TO SX60.
VECTOR SS=SS1 TO SS5.
COMPUTE XVAR=0.
LOOP #I=1 TO 60.
  COMPUTE #DIFF=SX(#I)/SW(#I) - XBAR.
  COMPUTE XVAR =XVAR + #DIFF * #DIFF.
END LOOP.
LOOP #I=1 TO 5.
COMPUTE SS(#I)=SS(#I)/SWT.
END LOOP.
COMPUTE SBAR=MEAN(SS1 TO SS5).
COMPUTE SVAR=VARIANCE(SS1 TO SS5).
COMPUTE XSE=SQRT(XVAR+(6/5)*SVAR).
PRINT FORMATS SBAR,SVAR,SVAR,XSE (F8.4).
LIST VARIABLES=BG12401,UWN,SWT,SBAR,XVAR,XSE.
```

Table 15-6. Standard error computation: Multiweight method using SAS with correction for imputation

```
DATA A;
      SET SYSFILE.MAI;
      ARRAY WT REPL01-REPL60;
      ARRAY WX WX1-WX60;
      ARRAY VALUE NRPP11-NRPP15;
      ARRAY WS WS1-WS5;
      IF (NRPP11 NE .);
      WTX=WEIGHT*NRPP11;
      DO OVER WT;
            WX=WT*NRPP11;
      END;
```

Table 15-6. Standard error computation: Multiweight method using SAS with correction for imputation — Continued

```
      DO OVER WS;
            WS=VALUE*WEIGHT;
      END;
PROC SUMMARY;
      CLASS BG12401;
      VAR WEIGHT REPL01-REPL60 WTX WX1-WX60 WS1-WS5;
      OUTPUT OUT=B N(WEIGHT)=UWN
            SUM(WEIGHT WTX REPL01-REPL60 WX1-WX60 WS1-WS5)=
              SWT SWX SW1-SW60 SX1-SX60 SS1-SS5;
DATA C;
      SET B;
      ARRAY SW SW1-SW60;
      ARRAY SX SX1-SX60;
      ARRAY SS SS1-SS5;
      XBAR=SWX/SWT;
      XVAR=0;
      DO OVER SW;
            DIFF=(SX/SW)-XBAR;
            XVAR=XVAR+DIFF*DIFF;
      END;
      DO OVER SS;
            SS=SS/SWT;
      END;
SBAR=MEAN(SS1,SS2,SS3,SS4,SS5);
      SVAR=VAR(SS1,SS2,SS3,SS4,SS5);
      XSE=SQRT(XVAR+(6/5)*SVAR);
PROC PRINT;
      VAR BG12401 UWN SWT SBAR XVAR SVAR XSE;
```

The flow of each program is as follows: for all variables the user instructs to be jackknifed, the program will generate two vectors of weighted sums and products. It will then sum these vectors across the entire file using the AGGREGATE (SPSS) or SUMMARY (SAS) procedure. Next, it will compute weighted means using the weighted sums; and finally, it will compute the estimated variance and standard error.

One advantage to this approach is that it will accomplish the computation in one pass of the data. Another advantage, afforded by the AGGREGATE (SPSS) and SUMMARY (SAS) procedures, is the facility to compute subgroup statistics by using the BREAK keyword (SPSS) or CLASS option (SAS) with the variable(s) defining the subgroups. All computations performed subsequent to the aggregation procedure are performed on each record of the collapsed file, corresponding to one of the subgroups. In the examples in Tables 15-3, 15-4, 15-5, and 15-6, the variable BG12401 (gender) is used as break control variable, and the derived statistics are printed for each gender code.

In Tables 15-3 and 15-4, the variable X may be any variable or transformation of variables except plausible values. In Tables 15-5 and 15-6, the vector or array named VALUE refers to one of the three sets of plausible values (prose, document, or quantitative). The prose plausible values are used in the examples.

The examples in Tables 15-3 to 15-6 apply to the household and non-incentive samples, which have 60 replicate weights. If the prison sample and the household sample are combined in one file and analyzed together, 15 additional replicate weights must be created for cases from the prison sample. Each of these weights (REPL46 – REPL60) should be set to the value of the final weight (WEIGHT) for each case. For just the prison sample, the coding should be changed to reflect that there are 45 replicate weights. The input data file for the household sample is SYSFILE.MAI, as used in the SAS examples, Tables 15-4 and 15-6. The input data files for the prison and non-incentive samples are SYSFILE.INC and SYSFILE.NON, respectively.

REFERENCES

Armstrong, J. 1975. "Monetary Incentives in Mail Surveys." *Public Opinion Quarterly*, 39, 111-116.

Baldwin, J., Kirsch, I. S., Rock, D., Yamamoto, K. 1995. *The Literacy Proficiencies of GED Examinees: Results from the GED-NALS Comparison Study*. Washington, DC: American Council of Education and Educational Testing Service.

Baxter, S., Nieva, V. F., Waksberg, J. 1984. *Improving Survey Response Rates: A Review of the Literature*. Rockville, MD: Westat, Inc.

Beaton, A. E. 1987. "Anchoring Scale Points." In A. E. Beaton et al. (Eds.), *Implementing the New Design: The NAEP 1983-84 Technical Report*, Section 10.5.2 (pp. 385-390). Princeton, NJ: Educational Testing Service.

Beaton, A. E., Allen, N. 1992. "Interpreting Scales Through Scale Anchoring." *Journal of Educational Statistics*, 17, 191-218.

Beaton, A. E., Johnson, B. G. 1990. "The Average Response Method of Scaling." *Journal of Educational Statistics*, 15, 9-38.

Berry, H., Kanouse, D. E. 1987. "Physician Response to a Mailed Survey." *Public Opinion Quarterly*, 51, 102-114.

Birnbaum, A. 1968. "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability." In F. M. Lord, M. R. Novick (Eds.), *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley Publishing.

Brackstone, G. B., Rao, J. N. K. 1979. "An Investigation of Raking Ratio Estimators." *Sankhya: The Indian Journal of Statistics*, 41(C-2), 97-144.

Brown, H., Prisuta, R., Jacobs, B., Campbell, A. 1996. *Literacy of Older Adults in America: Results from the National Adult Literacy Survey*. Washington, DC: U.S. Department of Education, National Center for Education Statistics (NCES 97-576).

Browne, M. W., Cudeck, R. 1993. "Alternative Ways of Assessing Model Fit." In K. A. Bollen, J. S. Long (Eds.), *Testing Structural Equation Models*. Sage Publications, in press.

Campbell, A., Kirsch, I. S., Kolstad, A. 1992. *Assessing Literacy: The Framework for the National Adult Literacy Survey*. Washington, DC: U.S. Department of Education, National Center for Education Statistics (NCES 93-113R).

Campbell, J. R., Donahue, P. L., Reese, C. M., Phillips, G. W. 1996. *NAEP 1994 Reading Report Card for the Nation and the States: Findings from the National Assessment of Educational Progress and Trial State Assessment*. Washington, DC: U.S. Department of Education, National Center for Education Statistics (Figure 6.4).

Cattell, R. B. 1966. "The Screen Test for the Number of Factors." *Multivariate Behavioral Research*, 1, 140-161.

Citro, C. F., Cohen, M. 1985. *The Bicentennial Census, New Directions for Methodology*. Washington, DC: National Academy Press.

Clark, H. H., & Haviland, S. E. 1977. "Comprehension and the Given-New Contract." In O. Freedle (Ed.), *Discourse Processes: Advances in Research and Theory, Vol. 1.* (pp. 1-40). Norwood, NJ: Ablex.

Cochran, W. G. 1977. *Sampling Techniques* (3rd edition). New York: John Wiley & Sons, Inc.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J. 1984. *Essentials of Psychological Testing* (4th edition). New York: Harper & Row.

Durbin, J. 1967. "Design of Multistage surveys for the Estimation of Sampling Errors." *Applied Statistics*, 16, 152-164.

Duncan, W. J. 1979. "Mail Questionnaires in Survey Research: A Review of Response Inducement Techniques." *Journal of Management*, 5(1), 39-55.

Fisher, D. 1981. "Functional Literacy Tests: A Model of Question-Answering and an Analysis of Errors." *Reading Research Quarterly*, 16, 18-8.

Fry, E B. December 1977. "Fry's Readability Graph: Clarifications, Validity, and Extension to Level 17." *Journal of Reading*, 21, 242-252.

Fry, E. B. 1981. "A Partial Reading Model Utilizing Language Unit Size by Frequency." In M. L. Kamil (Ed.), *Directions in Reading: Research and Instruction* (pp. 103-107). Washington, DC: National Reading Conference.

Fry, E. B. September 1975. "The Readability Principle." *Language Arts*, 52, 847-851

Godwin, R. 1979. "The Consequences of Large Monetary Incentives in Mail Surveys of Elites." *Public Opinion Quarterly*, 44, 378-387.

Groves, R. M., Couper, M. P. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons, Inc.

Gunn, J., Rhodes, I. N. 1981. "Physician Response Rates to a Telephone Survey: Effects of Monetary Incentive Level." *Public Opinion Quarterly*, 45, 109-115.

Guthrie, J. T., Siefert, M., Kirsch, I. S. 1986. "Effects of Education, Occupation, and Setting on Reading Practices." *American Educational Research Journal*, 23, 151-160.

Guttman, L. 1950. "The Basis for Scalogram Analysis." In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, J. A. Clausen (Eds.), *Studies in Social Psychology in World War II: Volume IV, Measurement and Prediction*, Chapter 3. Princeton, NJ: Princeton University Press.

Haigler, K. O., Harlow, C., O'Connor, P., Campbell, A. 1994. *Literacy Behind Prison Walls: Profiles of the Prison Population from the National Adult Literacy Survey*. Washington, DC: U.S. Department of Education, National Center for Education Statistics (NCES 94-102).

Huynh, H., (1998). "On score locations of binary and partial credit items and their applications to item mapping and criterion-referenced interpretations." *Journal of Educational and Behavioral Statistics*. 23, 37-58.

Ireland, C. T., Kullback, S. 1968. "Contingency Tables with Given Marginals." *Biometrika*, 55, 179-188.

Johnson, E. G., & Rust, K. F. 1992. "Population Inferences and Variance Estimation for NAEP Data." *Journal of Educational Statistics*, 17, 175-190.

Johnson, E. G. "Description of Percentages for Anchoring and Item Mapping" (unpublished internal Educational Testing Service memorandum, February 4, 1994).

Joreskog, K. G., Sorbom, D. 1979. *Advances In Factor Analysis And Structural Equation Models*. Cambridge, MA: Abt Books.

Joreskog, K. G., Sorbom, D. 1993. *LISREL8: Structural Equation Modeling with the Simplis Command Language*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kalton, G. 1981. *Compensation for Missing Survey Data*. Ann Arbor, MI: Institute for Social Research, University of Michigan.

Kanuk, L., & Berenson, C. 1975. "Mail Surveys and Response Rates: A Literature Review." *Journal of Marketing Research*, 12, 440-453.

Kirsch, I. S,. Jungeblut, A. et al. 1986. "Describing and Anchoring the Scales" and "Levels of Proficiency." In *Final Report: Literacy: Profiles of America's Young Adults*, Pp. III-9/III-10 and IV-11/IV-13, Princeton, NJ: Educational Testing Service.

Kirsch, I. S., Jungeblut, A. 1986a. *Literacy: Profiles of America's Young Adults*. Princeton, NJ: Educational Testing Service.

Kirsch, I. S., Jungeblut, A. 1986b. *Literacy: Profiles of America's Young Adults: Final Report*. Princeton, NJ: Educational Testing Service.

Kirsch, I. S., Mosenthal, P. B. 1990. "Exploring Document Literacy: Variables Underlying the Performance of Young Adults." *Reading Research Quarterly*, *25*(1), 5-30.

Kirsch, I. S., Mosenthal, P. B. December/January 1992b. "Integration Strategies: Higher-Order Thinking Applied to Documents." *Journal of Reading*, 36, 322-327.

Kirsch, I. S., Mosenthal, P. B. October 1992a. "How to Navigate a Document Using Locate Known/Need-to-Know Strategies." *Journal of Reading*, 36, 140-144.

Kirsch, I. S., Jungeblut, A., Campbell, A. 1992. *Beyond the School Doors: The Literacy Needs of Job Seekers Served by the U.S. Department of Labor*. Princeton, NJ: Educational Testing Service.

Kirsch, I. S., Jungeblut, A., Jenkins, L., Kolstad, A. 1993. *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, Washington, DC: Government Printing Office (NCES 93-275).

Kish, L. 1965. *Survey Sampling*. New York: John Wiley

Kolstad, A., Cohen, J., Baldi, S., Chan, T., deFur, E, Angeles, J. 1998. "Should NCES Adopt a Standard?" *The Response Probability Convention Used in Reporting Data From IRT Assessment Scales*. Washington, DC: American Institutes for Research.

Linsky, A. S. 1975. "Stimulating Responses to Mail Questionnaires: A Review." *Public Opinion Quarterly*, 39, 82-101.

Little, R. J. A., Rubin, D. 1983. "On Jointly Estimating Parameters and Missing Data." *American Statistician*, 37, 218-220.

Lord, F. M. 1953. "The Relation of Test Score to the Trait Underlying the Test." Reprinted in P. F. Lazarsfeld, N. W. Henry (Eds.), *Readings in Mathematical Social Science*. Cambridge, MA: Massachusetts Institute of Technology Press, 1966.

Lord, F. M. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Eribaum Associates.

Lord, F. M., & Novick, M. R. 1968. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.

Messick, S. 1989. "Validity." In R. L. Linn (Ed.), *Educational Measurement* (3rd edition). New York: Macmillan.

Messick, S., Beaton, A., Lord, F. 1983. *National Assessment of Educational Progress Reconsidered: A New Design for A New Era*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress (No. 83-1).

Meyer, B. J. F., Rice, G. E. 1984. "The Structure of Text." In P. D. Pearson, R. Barr, M. Kamil, P. Mosenthal (Eds.), *Handbook of Reading Research*, vol. 1, (pp. 319-352). White Plains, NY: Longman.

Meyer, B. J. F., Marsiske, M., Willis, S. L. 1993. "Text Processing Variables Predict the Readability of Everyday Documents Read by Older Adults." *Reading Research Quarterly*, 7, 400-419.

Mislevy, R. J. 1985. "Estimation of Latent Group Effects." *Journal of the American Statistical Association*, 80, 993-97.

Mislevy, R. J. 1990. Scaling procedures. In E. G. Johnson & R. Zwick (Eds.), *Focusing the New Design: The NAEP 1988 Technical Report*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress (No. 19-TR-20).

Mislevy, R. J. 1991. "Randomization-Based Inference About Latent Variables from Complex Samples." *Psychometrika*, 56, 177-196.

Mislevy, R. J. September 1984. "Estimating Latent Distributions." *Psychometrika*, 49, 359-81.

Mislevy, R. J., Bock, R. D. 1982. *Bilog: Item Analysis and Test Scoring with Binary Logistic Models* [Computer Program]. Morresville, IN: Scientific Software.

Mislevy, R. J., Sheehan, K. 1987. "Marginal Estimation Procedures." In A. H. Beaton (Ed.), *Implementing the New Design: The NAEP 1983-84 Technical Report*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress (No. 15-TR-20, pp. 293-360).

Mislevy, R. J., Beaton, A., Kaplan, B. A., Sheehan, K. Summer 1993. "Estimating Population Characteristics from Sparse Matrix Samples of Item Responses." *Journal of Educational Measurement*, 29, 133-161.

Mislevy, R. J., Johnson, E. G., Muraki, E. Summer 1993. "Scaling Procedures in the National Assessment of Educational Progress." *Journal of Educational Statistics*, 17, 131-154.

Mosenthal, P. B. 1985. "Defining the Expository Discourse Continuum: Towards a Taxonomy of Expository Text Types." *Poetics*, 14Y, 387-414.

Mosenthal, P. B., Kirsch, I. S. 1989. "Lists: The Building Blocks of Documents." *Journal of Reading*, 33, 58-60.

Mosenthal, P. B., Kirsch, I. S. 1991a. "Toward an Explanatory Model of Document Literacy." *Discourse Processes*, 14, 147-180.

Mosenthal, P. B., Kirsch, I. S. April 1993b. "Profiling Students' Document Strategy Abilities." *Journal of Reading*, 36, 578-583.

Mosenthal, P. B., Kirsch, I. S. February 1993a. "Generate Strategies: Coping Without Cues and Clues." *Journal of Reading*, 36, 416-419.

Mosenthal, P.B., Kirsch, I. S. May 1991b. "Information Types in Nonmimetic Documents: A Review of Biddle's Wipe-Clean Slate." *Journal of Reading*, 34, 654-660.

Mosenthal, P. B., Kirsch, I. S. November 1992. "Cycle Strategies in Document Search: from Here to There to Wherever." *Journal of Reading*, 36, 238-242.

Murphy, R. T. 1973. *Adult Functional Reading Study* (Final Report). Princeton, NJ: Educational Testing Service (PR-73-48).

Muthen, B. 1989. "Latent Variable Modeling in Heterogeneous Populations." *Psychometrika*, 54, 557-585.

Muthen, B., Khoo, S., Goff, S. 1994. *Multidimensional Description of Subgroup Differences in Mathematics Achievement Data From the 1992 National Assessment of Educational Progress.* Los Angeles, CA: UCLA, Graduate School of Education.

National Assessment of Educational Progress, Education Commission of the States. 1976. *Functional Literacy: Basic Reading Performance*. Denver, CO.

Novak, E., Seckman, B. S., & Stewart, R. D. 1977. "Motivations for Volunteering as Research Subjects." *Journal of Clinical Pharmacology*, 365-371.

Oh, H. L., Scheuren, F. 1978. "Some Unresolved Application Issues in Raking Ratio Estimation." *American Statistical Association Proceedings*, Section on Survey Research Methods, 723-728.

Phillips, G. W., Mullis, I. V. S., Bourque, M. L., Williams, P. L., Hambleton, R. K., Owen, E. H., Barton, P.E. 1993. *Interpreting NAEP Scales*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

SAS Institute, Inc. 1990. *SAS7 Language: Reference, Version 6, First Edition*. Cary, NC.

SAS Institute, Inc. 1991. Nonparametric Density Estimation. Section 3 4.2 in *SAS7 System for Statistical Graphics, First Edition*. Cary, NC.

SPSS, Inc. 1990. *SPSS7 Reference Guide*. Chicago.

Sticht, T. G., Armstrong, W. B. 1994. *Adult Literacy in the U.S.: A Compendium of Quantitative Data and Interpretive Comments*. Washington, DC: National Institute for Literacy, 103 and 116.

Sudman, S. 1976. *Applied Sampling*. New York: Academic Press.

Thomas, N. 1993. "Asymptotic Corrections for Multivariate Posterior Moments with Factored Likelihood Functions." *Journal of Computational and Graphical Statistics*, 2, 309-322.

Thorndike, R. L. 1982. *Applied Psychometrics*. Boston: Houghton Mifflin Co.

Train, G., Cahoon, L., Makens, P. 1978. "The Current Population Survey Variances, Interrelationships, and Design Effects." *American Statistical Association, Proceedings of the Section on Survey Research Methods*, 443-448.

Tucker, L. R., Lewis, C. 1973. "A Reliability Coefficient for Maximum Likelihood Factor Analysis." *Psychometrika*, 38, 1-10.

Tucker, L. R., Finkbeiner, C. T. 1981. *Transformation of Factors by Artificial Personal Probability Functions*. (ETS RR 82-88). Princeton, NJ: Educational Testing Service.

U.S. Department of Commerce, Census Bureau, Government Printing Office. 1992. *Alphabetical Index of Industries and Occupations*., Washington, DC.

U.S. Department of Justice, Bureau of Justice Statistics, Government Printing Office. 1991. *Survey Of Inmates Of State Correctional Facilities*. Washington, DC.

Walsh, T. C. 1972-73. "Selected Results from the 1972-73 Diary Surveys." *Journal of Marketing Research*, 14, 344-352.

Wingersky, M., Kaplan, B. A., Beaton, A. E. 1987. "Joint Estimation Procedures." In A. E. Beaton (Ed.), *Implementing the New Design: The NAEP 1983-84 Technical Report*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress (No. 15-TR-20, pp. 285-292).

Wolter, K. M. 1985. *Introduction to Variance Estimation*. New York: Springer-Verlag.

Yamamoto, K. Muraki, E., "Non-Linear Transformation of IRT Scale to Account for the Effect of Non-Normal Ability Distribution on Item Parameter Estimation" (paper presented at the 1991 Annual Meeting of the American Educational Research Association, Chicago, 1991).

Zusman, B. J., Duby. 1987. "An Evaluation of the Use of Monetary Incentives in Postsecondary Survey Research." *Journal of Research and Development in Education*, 20(4), 73-78.

Appendix A

ESTIMATED ITEM PARAMETERS

Appendix A: Estimated Item Parameters
NATIONAL ADULT LITERACY SURVEY

PROSE								
NAL ID	A	SE	B	SE	C	SE	BLOCK	ITEM
AB21101	1.12541	0.04204	-1.90129	0.04888	0.00000	0.00000	B	11
AB21201	1.07025	0.02913	-1.12479	0.02744	0.00000	0.00000	B	12
AB30501	0.59088	0.01569	0.59334	0.02292	0.00000	0.00000	C	5
AB30601	0.91549	0.02386	0.34748	0.01890	0.00000	0.00000	C	6
AB31201	0.72520	0.01805	0.69108	0.02013	0.00000	0.00000	C	12
AB40901	0.82634	0.01892	0.16578	0.01763	0.00000	0.00000	D	9
AB41001	0.62236	0.02095	-1.43344	0.05300	0.00000	0.00000	D	10
AB50101	0.46634	0.01641	2.11237	0.05732	0.00000	0.00000	E	1
AB50201	1.16067	0.03631	0.86105	0.01786	0.19631	0.00000	E	2
AB60201	1.24079	0.02754	-0.44014	0.01576	0.00000	0.00000	F	2
AB60601	0.80877	0.01990	-0.31995	0.02154	0.00000	0.00000	F	6
AB70401	0.70551	0.01851	-0.76529	0.02921	0.00000	0.00000	G	4
AB71001	1.37774	0.04293	-0.30561	0.02074	0.26646	0.00000	G	10
AB71101	0.78267	0.02175	0.48208	0.02147	0.00000	0.00000	G	11
NC00301	0.89266	0.03490	-3.22823	0.09046	0.00000	0.00000	CORE	3
NC00401	0.76513	0.01644	-1.93599	0.03498	0.00000	0.00000	CORE	4
N010101	0.86858	0.02570	0.60760	0.02239	0.00000	0.00000	A	1
N010201	1.05942	0.03120	-0.40290	0.02249	0.00000	0.00000	A	2
N010301	0.78697	0.03141	2.13821	0.05369	0.00000	0.00000	A	3
N080101	1.32888	0.05177	-1.44713	0.03620	0.00000	0.00000	H	1
N080201	1.51648	0.04358	-0.38940	0.01722	0.00000	0.00000	H	2
N080301	0.61896	0.02120	0.48631	0.02867	0.00000	0.00000	H	3
N090601	1.87803	0.06406	-0.74821	0.01880	0.00000	0.00000	I	6
N090701	1.80488	0.06026	-0.69902	0.01881	0.00000	0.00000	I	7
N090801	1.23905	0.03753	1.09108	0.02071	0.00000	0.00000	I	8
N100101	1.46688	0.05219	-1.14666	0.02728	0.00000	0.00000	J	1
N100201	1.29706	0.03719	-0.34594	0.01889	0.00000	0.00000	J	2
N100301	1.18710	0.03428	-0.34391	0.02013	0.00000	0.00000	J	3
N100401	0.84168	0.02718	1.23642	0.02927	0.00000	0.00000	J	4

PROSE								
NAL ID	**A**	**SE**	**B**	**SE**	**C**	**SE**	**BLOCK**	**ITEM**
N110101	0.98825	0.03283	-0.97143	0.03235	0.00000	0.00000	K	1
N110401	0.77035	0.02454	-0.19160	0.02722	0.00000	0.00000	K	4
N110501	0.93935	0.03062	-0.73081	0.03031	0.00000	0.00000	K	5
N110601	1.04481	0.03940	1.95448	0.03872	0.00000	0.00000	K	6
N120301	1.07451	0.03037	0.14152	0.01909	0.00000	0.00000	L	3
N120401	1.16237	0.03232	0.22911	0.01789	0.00000	0.00000	L	4
N120501	0.92685	0.03702	2.10722	0.04855	0.00000	0.00000	L	5
N120901	0.88878	0.04458	-2.06119	0.08039	0.00000	0.00000	L	9
N130201	1.08905	0.03043	0.31566	0.01866	0.00000	0.00000	M	2
N130301	0.97807	0.03057	1.21360	0.02599	0.00000	0.00000	M	3
N130401	1.57618	0.04556	0.97858	0.01654	0.00000	0.00000	M	4
N130801	0.73534	0.02758	-1.01262	0.04397	0.00000	0.00000	M	8

DOCUMENT								
NAL ID	A	SE	B	SE	C	SE	BLOCK	ITEM
SCOR100	0.50477	0.02590	-4.80384	0.24886	0.00000	0.00000	CORE	1
SCOR300	0.91752	0.02517	-2.52522	0.05845	0.00000	0.00000	CORE	2
AB20101	1.15374	0.04514	-0.19348	0.05430	0.22802	0.03016	B	1
AB20201	0.93557	0.03047	-0.02307	0.04597	0.09655	0.02347	B	2
AB20301	1.08976	0.03692	0.68447	0.03114	0.14215	0.01513	B	3
AB20401	0.47887	0.01971	-0.46797	0.11156	0.14409	0.03626	B	4
AB20501	0.41467	0.01723	-0.77194	0.11128	0.08838	0.03129	B	5
AB20601	1.07779	0.03477	-0.14302	0.04175	0.10587	0.02339	B	6
AB20701	0.52151	0.02476	0.29345	0.10693	0.13091	0.03537	B	7
AB20801	1.28204	0.04413	0.90155	0.02487	0.14404	0.01259	B	8
AB20901	1.16899	0.03276	1.52080	0.02177	0.16275	0.00826	B	9
AB21001	0.73014	0.03145	0.52075	0.06625	0.14440	0.02693	B	10
AB21501	0.79908	0.02419	-0.57187	0.03812	0.00000	0.00000	B	15
AB30101	0.95382	0.02782	-0.95612	0.03603	0.00000	0.00000	C	1
AB30301	0.90436	0.02562	-0.84497	0.03478	0.00000	0.00000	C	3
AB30401	0.66593	0.01762	-0.08909	0.02804	0.00000	0.00000	C	4
AB30701	0.96073	0.02640	-0.70250	0.03090	0.00000	0.00000	C	7
AB30801	0.70440	0.01747	0.92937	0.01961	0.00000	0.00000	C	8
AB30901	0.29941	0.01333	-0.00016	0.05623	0.00000	0.00000	C	9
AB31001	0.83122	0.01999	0.28507	0.02026	0.00000	0.00000	C	10
AB31101	0.76160	0.02031	-0.25692	0.02883	0.00000	0.00000	C	11
AB31301	0.72146	0.02463	-1.17011	0.05551	0.00000	0.00000	C	13
AB40101	0.82062	0.02471	-1.06345	0.04192	0.00000	0.00000	D	1
AB40401	1.10828	0.02302	0.71717	0.01385	0.00000	0.00000	D	4
AB50401	0.77263	0.02232	-0.88270	0.03944	0.00000	0.00000	E	4
AB50402	0.77114	0.01825	0.39658	0.01995	0.00000	0.00000	E	4
AB50501	0.35970	0.01410	-0.51177	0.06071	0.00000	0.00000	E	5
AB50601	1.00104	0.02392	-0.08331	0.02093	0.00000	0.00000	E	6
AB50701	1.18249	0.02992	-0.37345	0.02205	0.00000	0.00000	E	7
AB50801	0.73380	0.02581	-1.36581	0.06034	0.00000	0.00000	E	8
AB50901	0.88412	0.02266	-0.19958	0.02533	0.00000	0.00000	E	9
AB60101	1.25423	0.03179	-0.49715	0.02132	0.00000	0.00000	F	1

DOCUMENT								
NAL ID	**A**	**SE**	**B**	**SE**	**C**	**SE**	**BLOCK**	**ITEM**
AB60102	1.40817	0.03524	-0.42516	0.01874	0.00000	0.00000	F	1
AB60103	0.99335	0.02654	-0.67415	0.02842	0.00000	0.00000	F	1
AB60104	1.53757	0.04048	-0.52476	0.01882	0.00000	0.00000	F	1
AB60301	1.45438	0.05419	-1.28304	0.03604	0.00000	0.00000	F	3
AB60302	1.06892	0.03859	-1.43441	0.04847	0.00000	0.00000	F	3
AB60303	0.90384	0.02447	-0.68004	0.03076	0.00000	0.00000	F	3
AB60304	0.89514	0.01998	0.46187	0.01722	0.00000	0.00000	F	3
AB60305	0.64035	0.01743	-0.22092	0.03069	0.00000	0.00000	F	3
AB60306	0.94756	0.02705	-0.86793	0.03380	0.00000	0.00000	F	3
AB60501	1.10267	0.02487	1.93775	0.01926	0.00000	0.00000	F	5
AB60502	1.08158	0.02316	0.78254	0.01425	0.00000	0.00000	F	5
AB60701	1.17906	0.04573	-1.29572	0.04760	0.00000	0.00000	F	7
AB60801	1.01657	0.04270	-1.53935	0.06363	0.00000	0.00000	F	8
AB60802	1.23196	0.05855	-1.62019	0.06426	0.00000	0.00000	F	8
AB60803	1.43895	0.07608	-1.64959	0.06318	0.00000	0.00000	F	8
AB61001	0.76644	0.03042	-1.45425	0.06986	0.00000	0.00000	F	10
AB70104	0.54278	0.02446	-2.33730	0.11924	0.00000	0.00000	G	1
AB70301	0.75388	0.01890	-0.13404	0.02574	0.00000	0.00000	G	3
AB70701	0.79927	0.01974	-0.12643	0.02461	0.00000	0.00000	G	7
AB70801	1.14286	0.03314	-0.88087	0.03047	0.00000	0.00000	G	8
AB70901	0.99034	0.03059	-1.08894	0.03902	0.00000	0.00000	G	9
AB70902	0.85846	0.02167	-0.30332	0.02599	0.00000	0.00000	G	9
AB70903	1.26606	0.03805	-0.92160	0.02946	0.00000	0.00000	G	9
AB71201	0.93964	0.04169	-1.80154	0.07770	0.00000	0.00000	G	12
N010401	0.90290	0.03889	-1.34078	0.06201	0.00000	0.00000	A	4
N010801	0.80799	0.02813	-0.46332	0.03855	0.00000	0.00000	A	8
N010901	0.72077	0.02450	1.70210	0.03226	0.00000	0.00000	A	9
N011001	0.64599	0.02247	0.26092	0.03246	0.00000	0.00000	A	10
AB70105	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	G	1
N080601	1.03992	0.02930	0.50524	0.02007	0.00000	0.00000	H	6
N080701	1.09493	0.03435	-0.31250	0.02734	0.00000	0.00000	H	7
N080801	0.76325	0.02370	0.56953	0.02519	0.00000	0.00000	H	8

DOCUMENT								
NAL ID	A	SE	B	SE	C	SE	BLOCK	ITEM
N080802	1.35711	0.04834	-0.68388	0.02975	0.00000	0.00000	H	8
N090301	1.12350	0.04897	-1.22466	0.05147	0.00000	0.00000	I	3
N090401	0.98761	0.03333	-0.44820	0.03256	0.00000	0.00000	I	4
N090501	0.67120	0.02441	-0.30100	0.04074	0.00000	0.00000	I	5
N100501	1.03859	0.02982	0.48633	0.02010	0.00000	0.00000	J	5
N100601	1.13463	0.03235	1.28460	0.01915	0.00000	0.00000	J	6
N100701	1.12769	0.03464	2.30042	0.03207	0.00000	0.00000	J	7
N110301	0.81191	0.02964	-0.74287	0.04522	0.00000	0.00000	K	3
N110302	0.71414	0.02863	-1.02576	0.05940	0.00000	0.00000	K	3
N110701	0.46961	0.02024	0.12592	0.04748	0.00000	0.00000	K	7
N110901	0.82921	0.03133	1.88254	0.03287	0.00000	0.00000	K	9
N120101	0.98552	0.03638	-0.80163	0.04050	0.00000	0.00000	L	1
N120201	0.84221	0.02873	-0.40312	0.03522	0.00000	0.00000	L	2
N120601	0.79549	0.03735	-1.48811	0.07739	0.00000	0.00000	L	6
N130101	1.61922	0.04936	-0.09574	0.01793	0.00000	0.00000	M	1
N130102	1.27041	0.04351	-0.54440	0.02829	0.00000	0.00000	M	1
N130103	2.10560	0.07196	-0.29053	0.01687	0.00000	0.00000	M	1
N130104	2.15938	0.06948	-0.11128	0.01494	0.00000	0.00000	M	1

QUANTITATIVE								
NAL ID	A	SE	B	SE	C	SE	BLOCK	ITEM
AB40201	0.81800	0.01920	0.45529	0.01728	0.00000	0.00000	D	2
AB40301	0.81591	0.03463	0.21661	0.02953	0.44751	0.00000	D	3
AB40501	0.90966	0.02044	0.00576	0.01605	0.00000	0.00000	D	5
AB40601	0.95267	0.02175	-0.52245	0.01879	0.00000	0.00000	D	6
AB40701	1.59726	0.03487	-0.50086	0.01312	0.00000	0.00000	D	7
AB40702	1.93608	0.04230	-0.34473	0.01086	0.00000	0.00000	D	7
AB40703	1.87367	0.04067	-0.33187	0.01101	0.00000	0.00000	D	7
AB40704	1.97024	0.04291	-0.29467	0.01051	0.00000	0.00000	D	7
AB50301	0.60154	0.02076	1.52255	0.04321	0.00000	0.00000	E	3
AB50403	0.60917	0.01605	0.60064	0.02315	0.00000	0.00000	E	4
AB50404	0.96847	0.02339	-0.95191	0.02270	0.00000	0.00000	E	4
AB60901	0.50410	0.01569	-0.35528	0.03111	0.00000	0.00000	F	9
AB70501	0.89357	0.01985	0.09082	0.01626	0.00000	0.00000	G	5
AB70601	0.87252	0.01961	0.38444	0.01658	0.00000	0.00000	G	6
AB70904	0.86946	0.02907	-1.97025	0.04982	0.00000	0.00000	G	9
NC00501	0.66137	0.01774	-2.79229	0.06017	0.00000	0.00000	CORE	5
NC00601	0.71757	0.01389	-1.69034	0.02838	0.00000	0.00000	CORE	6
N010501	0.85146	0.02699	-0.76868	0.02957	0.00000	0.00000	A	5
N010601	1.12179	0.03237	0.71700	0.01979	0.00000	0.00000	A	6
N010701	1.03353	0.02945	0.41126	0.01952	0.00000	0.00000	A	7
N011101	1.03451	0.03042	0.19559	0.01995	0.00000	0.00000	A	11
N080401	0.69649	0.02275	0.52081	0.02695	0.00000	0.00000	H	4
N080501	0.75721	0.02374	-0.24795	0.02594	0.00000	0.00000	H	5
N080901	0.85076	0.02710	0.85616	0.02603	0.00000	0.00000	H	9
N081001	0.73272	0.02392	0.12226	0.02560	0.00000	0.00000	H	10
N090101	1.34693	0.03733	-0.01808	0.01640	0.00000	0.00000	I	1
N090201	1.67703	0.04789	-0.34944	0.01554	0.00000	0.00000	I	2
N090901	0.78973	0.02875	-1.00384	0.04000	0.00000	0.00000	I	9
N091001	0.63476	0.02629	1.37112	0.04532	0.00000	0.00000	I	10
N100801	0.64782	0.02784	-1.73791	0.06814	0.00000	0.00000	J	8
N100901	0.62252	0.02245	-0.26352	0.03224	0.00000	0.00000	J	9
N101001	0.94371	0.03186	-0.83751	0.03112	0.00000	0.00000	J	10

QUANTITATIVE								
NAL ID	A	SE	B	SE	C	SE	BLOCK	ITEM
N110201	1.03343	0.03035	0.74076	0.02128	0.00000	0.00000	K	2
N110303	0.78964	0.03160	-1.73049	0.05686	0.00000	0.00000	K	3
N110801	0.88151	0.02997	-0.49499	0.02909	0.00000	0.00000	K	8
N120701	0.90911	0.02939	-0.84546	0.02993	0.00000	0.00000	L	7
N120801	1.01300	0.03078	0.83053	0.02265	0.00000	0.00000	L	8
N121001	1.01768	0.03123	0.21765	0.02054	0.00000	0.00000	L	10
N121101	0.95960	0.03595	1.51759	0.03591	0.00000	0.00000	L	11
N130501	0.65539	0.02303	0.81932	0.03207	0.00000	0.00000	M	5
N130601	1.11183	0.03182	-0.18439	0.01963	0.00000	0.00000	M	6
N130701	0.84576	0.03479	1.96223	0.05280	0.00000	0.00000	M	7
N130901	0.94575	0.03054	0.49942	0.02255	0.00000	0.00000	M	9

Appendix B

CONDITIONING VARIABLES

Appendix B: Conditioning Variables
NATIONAL ADULT LITERACY SURVEY
Contrast Coding for the MGROUP Conditioning Variables

	Item #	Question	Contrast Coding	Variable Coding
	B-1	Highest level of education completed in the USA?	1000000000000	Still in high school
			0100000000000	Less than high school
			0010000000000	Some high school
			0001000000000	GED or high school equivalency
			0000000000000	High school graduate
			0000100000000	Vocational
			0000010000000	College: less than two years
			0000001000000	College: Associate's degree
			0000000100000	College: two years or more, no degree
			0000000010000	College graduate (B.S. or B.A.)
			0000000001000	Postgraduate, no degree
			0000000000100	Postgraduate, degree
			0000000000010	Don't know
			0000000000001	Missing
	B-3	Studied for a GED or high school equivalency?	00	Yes
			10	No
			01	Missing
SECTION F: Demographic Information				
	F-1	Highest level of education your mother completed?	100000000000	Less than high school
			010000000000	Some high school
			001000000000	GED or high school equivalency
			000000000000	High school graduate
			000100000000	Vocational
			000010000000	College: less than two years
			000001000000	College: Associate's degree
			000000100000	College: two years or more, no degree
			000000010000	College graduate (B.S. or B.A.)
			000000001000	Postgraduate, no degree
			000000000100	Postgraduate, degree
			000000000010	Don't know
			000000000001	Missing
	F-2	Highest level of education your father completed?	100000000000	Less than high school
			010000000000	Some high school
			001000000000	GED or high school equivalency
			000000000000	High school graduate
			000100000000	Vocational
			000010000000	College: less than two years
			000001000000	College: Associate's degree
			000000100000	College: two years or more, no degree
			000000010000	College graduate (B.S. or B.A.)
			000000001000	Postgraduate, no degree
			000000000100	Postgraduate, degree
			000000000010	Don't know
			000000000001	Missing
	F-3	Current marital status?	100000	Never married
			000000	Married, living with spouse
			010000	Married, spouse living elsewhere
			001000	Living as married
			000100	Separated or divorced
			000010	Widowed
			000001	Missing
	F-4	Including yourself, how many in family employed part-time?	0000	None
			1000	One
			0100	Two
			0010	Three or more
			0001	Missing

	Item #	Question	Contrast Coding	Variable Coding
	F-5	Including yourself, how many in family employed full-time?	1000 0000 0100 0010 0001	None One Two Three or more Missing
	F-6	Did you or a family member receive any of the following during the past 12 months? Social Security or Railroad Retirement payments?	10 00 01	Yes No Missing
		Supplemental Security income?	10 00 01	Yes No Missing
		Other retirement, survivor, or disability payments?	10 00 01	Yes No Missing
		Aid to Families with Dependent Children?	10 00 01	Yes No Missing
		Food stamps?	10 00 01	Yes No Missing
		Interest from savings or other bank accounts?	10 00 01	Yes No Missing
		Dividend income from stocks, mutual funds, rental property, etc?	10 00 01	Yes No Missing
		Income from other sources?	10 00 01	Yes No Missing
		None.	10 00 01	Yes No Missing
	F-7	Total family income?	1000000000 0100000000 0010000000 0001000000 0000100000 0000000000 0000010000 0000001000 0000000100 0000000010 0000000001 0000000001 0000000001	Under $5,000 $ 5,000 - $9,999 $10,000 - $14,999 $15,000 - $19,999 $20,000 - $29,999 $30,000 - $39,999 $40,000 - $49,999 $50,000 - $74,999 $75,000 - $99,999 $100,000 and over Refused Don't know Missing
	F-8	Personal income?	10000000000 01000000000 00100000000 00010000000 00001000000 00000000000 00000100000 00000010000 00000001000 00000000100 00000000010 00000000001 00000000001 00000000001	Under $5,000 $ 5,000 - $9,999 $10,000 - $14,999 $15,000 - $19,999 $20,000 - $29,999 $30,000 - $39,999 $40,000 - $49,999 $50,000 - $74,999 $75,000 - $99,999 $100,000 and over No personal income Refused Don't know Missing

	Item #	Question	Contrast Coding	Variable Coding
	F-9	Race/ethnicity?	1000000000	Black not Hispanic
			0100000000	Hispanic, Mexicano, Mexican, Mexican American, or Chicano
			0010000000	Hispanic, Puerto Rican
			0001000000	Hispanic, Cuban
			0000100000	Hispanic, Central/South American or other Spanish
			0000010000	Asian
			0000001000	American Indian or Alaskan Native
			0000000100	Pacific Islander
			0000000000	White
			0000000010	Other
			0000000001	Missing
	F-10	Are you of Spanish or Hispanic origin or descent?	10	Yes
			00	No
			01	Missing
	F-11	Which best describes your Hispanic origin?	00000	Mexicano, Mexican, Mexican American, or Chicano
			10000	Puerto Rican
			01000	Cuban
			00100	Central/South American
			00010	Other Spanish/Hispanic
			00001	Missing or does not apply
	F-12	Date of birth-Age?		Actual age (Range 16-100) Missing = 40 (9 cases)
	F-13	Sex?	10	Male
			00	Female
			01	Missing
		Census Region of the country. (cannot be missing)	100	Northeast
			000	Midwest
			010	South
			001	West
SECTION A: General and Language Background				
Derived	A-1	In what country were you born?	00000	USA
			10000	Spanish speaking countries
			01000	European language countries
			00100	Asian language countries
			00010	All other counties
			00001	Missing
	A-2	How many years have you lived in the United States?	1	1 - 5 years
			2	6 - 10 years
			3	11 - 15 years
			4	16 - 20 years
			5	21 - 30 years
			6	31 - 40 years
			7	41 - 50 years
			8	51 or more years
			9	Missing by design - skip pattern
Derived	A-2b	Age of arrival in the USA ? (Age minus median of A-2)	0	Born in the USA or missing
			0	Negative values
				Age 1 to 90
	A-3	Highest education before coming to the USA?	100	Born in the USA or missing
			000	Primary or elementary (codes 2 & 3)
			010	Some schooling (codes 4 to 7)
			001	No schooling (code 1)

	Item #	Question	Contrast Coding	Variable Coding
Derived	A-4	Languages spoken in the home while growing up?	0000000000	English only
			1000000000	English & Spanish
			0100000000	English & European
			0010000000	English & Asian
			0001000000	Spanish & other
			0000100000	Other & other
			0000010000	Spanish only
			0000001000	European only
			0000000100	Asian only
			0000000010	English & other
			0000000001	Missing
Derived	A-5	Languages learned to speak before starting school?	0000000000	English only
			1000000000	English & Spanish
			0100000000	English & European
			0010000000	English & Asian
			0001000000	Spanish & other
			0000100000	Other & other
			0000010000	Spanish only
			0000001000	European only
			0000000100	Asian only
			0000000010	English & other
			0000000001	Missing
Derived	A-6	What language did you first learn to read and write?	000000	English
			100000	Spanish
			010000	European
			001000	Asian
			000100	Other non-English language
			000010	Never learned to read and write
			000001	Missing
	A-7	How old were you when you learned to speak English?	000000	1 - 4 years old
			100000	5 - 10 years old
			010000	11 - 15 years old
			001000	16 - 20 years old
			000100	21 or older
			000010	Does not speak English
			000001	Missing
	A-8	With regard to (NON-ENGLISH LANGUAGE IN A-5) how well do you understand it when it is spoken to you?	1000	Not at all
			0100	Not well
			0000	Well
			0010	Very well
			0001	Missing
		speak it?	1000	Not at all
			0100	Not well
			0000	Well
			0010	Very well
			0001	Missing
		read it?	1000	Not at all
			0100	Not well
			0000	Well
			0010	Very well
			0001	Missing
		write it?	1000	Not at all
			0100	Not well
			0000	Well
			0010	Very well
			0001	Missing
	A-9	With regard to (NON-ENGLISH LANGUAGE IN A-5) how often do you listen to radio programs in that language?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing

	Item #	Question	Contrast Coding	Variable Coding
	A-9 (contd.)	watch television programs in that language?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		read newspapers, magazines or books in that language?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		write or fill out letters or forms in that language?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	A-10	Language used in each of the following situations	000000	Always English
			100000	More English than other language
			010000	English and other language equally
		At home?	001000	More other language than English
			000100	Always other language
			000010	Does not apply
			000001	Missing
		At work?	000000	Always English
			100000	More English than other language
			010000	English and other language equally
			001000	More other language than English
			000100	Always other language
			000010	Does not apply
			000001	Missing
		While shopping in your neighborhood?	000000	Always English
			100000	More English than other language
			010000	English and other language equally
			001000	More other language than English
			000100	Always other language
			000010	Does not apply
			000001	Missing
		When visiting relatives or friends?	000000	Always English
			100000	More English than other language
			010000	English and other language equally
			001000	More other language than English
			000100	Always other language
			000010	Does not apply
			000001	Missing
	A-11	Have you taken a course to learn how to read and write English (2nd lang)?	00	Yes
			10	No
			01	Missing
	A-12	Did you complete this course?	00	Yes
			10	No
			01	Missing
	A-13	Have you taken a course to learn to speak or understand English (2nd lang)?	00	Yes
			10	No
			01	Missing
	A-14	Did you complete this course?	00	Yes
			10	No
			01	Missing
Derived	A-15	Which language do you usually speak now?	00000	English
			10000	Spanish
			01000	European
			00100	Asian
			00010	Other Non-English language
			00001	Missing

	Item #	Question	Contrast Coding	Variable Coding
Derived	A-16	What other language do you often speak now?	000000	English
			100000	Spanish
			010000	European
			001000	Asian
			000100	Other non-English language
			000010	No other language spoken
			000001	Missing
	A-17	With regard to the English language, how well do you understand it when it is spoken to you?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		speak it?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		read it?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		write it?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		do arithmetic problems?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
SECTION B: Educational Background and Experiences				
	B-2	Main reason you stopped your schooling?	100000	Financial problems
			010000	Went to work or into the military
			001000	Pregnancy
			000000	Lost interest or behavior problems, and academic problems
			000100	Family or personal problems
			000010	Other
			000001	Missing
	B-4	Kind of high school program?	1000	General
			0000	Vocational, technical, or trade
			0100	College preparatory
			0010	The same for everyone
			0001	Missing
	B-5	Currently enrolled in school or college, full- or part-time?	00	Yes
			10	No
			01	Missing
	B-6	Upon completing current schooling what degree or certificate do you expect to earn?	1000000	None
			0100000	Other
			0000000	High school diploma or equivalency
			0010000	Vocational, trade, or business
			0001000	Two-year college degree
			0000100	Four- or five-year college degree (B.S., B.A.)
			0000010	Master's, Ph.D., M.D., or other advanced degree
			0000001	Missing
	B-7	Enrolled now or have taken a program to improve basic skills?	00	Yes
			10	No
			01	Missing

	Item #	Question	Contrast Coding	Variable Coding
	B-8	Which program? Training program or course given by employer or union?	000	Yes
			100	No
			010	Don't know
			001	Missing
		Publicly sponsored training program such as JTPA or ABE?	000	Yes
			100	No
			010	Don't know
			001	Missing
		Tutoring program sponsored by library, church, or community?	000	Yes
			100	No
			010	Don't know
			001	Missing
		Any other program?	000	Yes
			100	No
			010	Don't know
			001	Missing
	B-9	If 'Yes' to B-8 Employer/union training program?	1000	Still enrolled
			0000	Within the past year
			0100	1 - 5 years ago
			0010	More than five years ago
			0001	Missing
		Publicly sponsored program, JTPA or ABE?	1000	Still enrolled
			0000	Within the past year
			0100	1 - 5 years ago
			0010	More than five years ago
			0001	Missing
		Tutoring program by library, church or community?	1000	Still enrolled
			0000	Within the past year
			0100	1 - 5 years ago
			0010	More than five years ago
			0001	Missing
		Any other program, military, prisons or institutions?	1000	Still enrolled
			0000	Within the past year
			0100	1 - 5 years ago
			0010	More than five years ago
			0001	Missing
	B-10	Where did you primarily learn to read newspapers, magazines, or books?	10000	Mostly in school
			00000	Mostly at home or in the community
			01000	Mostly at work
			00100	Other
			00010	Did not learn
			00001	Missing
		read graphs, diagrams, or maps?	10000	Mostly in school
			00000	Mostly at home or in the community
			01000	Mostly at work
			00100	Other
			00010	Did not learn
			00001	Missing
		fill out forms?	10000	Mostly in school
			00000	Mostly at home or in the community
			01000	Mostly at work
			00100	Other
			00010	Did not learn
			00001	Missing
		write letters, notes, memos, or reports?	10000	Mostly in school
			00000	Mostly at home or in the community
			01000	Mostly at work
			00100	Other
			00010	Did not learn
			00001	Missing
	B-11	Physical, mental or other health condition that stops your participation fully in work, school, housework or other?	00	No
			10	Yes
			01	Missing

	Item #	Question	Contrast Coding	Variable Coding
	B-12	Difficulty seeing words and letters in newspaper even when wearing glasses or contact lenses, if you usually wear them?	00 10 01	No Yes Missing
	B-13	Difficulty hearing a normal conversation even when using a hearing aid, if you usually wear one?	00 10 01	No Yes Missing
	B-14	Do you currently have any of these conditions? A learning disability?	00 10 01	No Yes Missing
		Any mental or emotional condition?	00 10 01	No Yes Missing
		Mental retardation?	00 10 01	No Yes Missing
		A speech disability?	00 10 01	No Yes Missing
		A physical disability?	00 10 01	No Yes Missing
		A long-term illness (6 months or more)?	00 10 01	No Yes Missing
		Any other health impairment?	00 10 01	No Yes Missing
SECTION C: Political and Social Participation				
	C-1	I'd like to find out how you usually get information about current events, public affairs and the government. newspapers?	1000 0100 0000 0010 0001	None A little Some A lot Missing
		magazines?	1000 0100 0000 0010 0001	None A little Some A lot Missing
		radio?	1000 0100 0000 0010 0001	None A little Some A lot Missing
		television?	1000 100 0000 0010 0001	None A little Some A lot Missing
		family members, friends, or coworkers?	1000 0100 0000 0010 0001	None A little Some A lot Missing
	C-2	How many hours do you usually watch television each day?	1000000 0100000 0010000 0000000 0001000 0000100 0000010 0000001	None 1 hour or less 2 hours 3 hours 4 hours 5 hours 6 or more hours Missing

	Item #	Question	Contrast Coding	Variable Coding
	C-3	How often do you use the services of a library?	10000 01000 00100 00000 00010 00001	Daily Weekly Monthly Once or twice a year Never Missing
	C-4	Have you voted in a national or state election in the USA within the past five years?	000 100 010 001	Yes No Not eligible to vote Missing
SECTION D: Labor Force Participation				
	D-1	Questions about what you were doing last week. Working a full-time job for pay or profit, 35 hours or more?	10 00 01	Yes No Missing
		Working for pay or profit part-time, 1 to 34 hours?	10 00 01	Yes No Missing
		Working two or more part-time jobs, totaling 35 or more hours?	10 00 01	Yes No Missing
		Unemployed, laid off, or looking for work?	10 00 01	Yes No Missing
		Have a job but temporary illness, vacation, or work stoppage?	10 00 01	Yes No Missing
		Have a job but on family leave (maternity or paternity leave)?	10 00 01	Yes No Missing
		In school?	10 00 01	Yes No Missing
		Keeping house?	10 00 01	Yes No Missing
		Retired?	10 00 01	Yes No Missing
		Doing volunteer work?	10 00 01	Yes No Missing
		Other?	10 00 01	Yes No Missing
	D-2	Have you looked for a job during the past four weeks?	10 00 01	Yes No Missing
	D-3	Last week, what was your total weekly wage or salary?	0	7's, 8's, 9's and missing Used actual weekly wage if specified
	D-5	Number of weeks worked for pay during the last 12 months?	0 0	None 88, 99 and missing Used actual number of weeks if specified
	D-6	Of the weeks you were not employed, what were you doing? Looking for work?	10 00 01	Yes No Missing

	Item #	Question	Contrast Coding	Variable Coding
	D-6 (contd.)	In school or training?	10	Yes
			00	No
			01	Missing
		Keeping house?	10	Yes
			00	No
			01	Missing
		Retired?	10	Yes
			00	No
			01	Missing
		Volunteering?	10	Yes
			00	No
			01	Missing
		Doing something else?	10	Yes
			00	No
			01	Missing
	D-7	For the past 12 months, what was your average weekly wage?	0	7's, 8's, 9's and missing Used actual average weekly wage, if specified
	D-9	Which of the following describes your work history?	100	Held a paying job within the last three years
			000	Held a paying job but not within the last three years
			010	Never been employed for pay full-time or part-time
			001	Missing

D-10 What kind of business or industry (do/did) you work in?

10000000000000000000000	Agriculture, forestry, fishing
01000000000000000000000	Mining
00100000000000000000000	Construction
00010000000000000000000	Manufacturing - Durable
00001000000000000000000	Manufacturing - Nondurable
00000100000000000000000	Transportation
00000010000000000000000	Communications
00000001000000000000000	Utilities and sanitary services
00000000100000000000000	Wholesale trade
00000000010000000000000	Retail trade - Eating and drinking places
00000000001000000000000	Retail trade - Grocery, food, dairy, and bakery stores
00000000000000000000000	Retail trade - All other retail trade
00000000000100000000000	Finance - Banking and stock exchanges or commodities
00000000000010000000000	Finance - Insurance
00000000000001000000000	Finance - Real estate
00000000000000100000000	Services - Business and repair
00000000000000010000000	Services - Personal
00000000000000001000000	Services - Entertainment and recreation
00000000000000000100000	Services - Professional and related
00000000000000000010000	Public Administration
00000000000000000001000	Military
00000000000000000000100	Missing
00000000000000000000010	Not worked in past 3 years
00000000000000000000001	Never worked

D-11 What (is/was) your current occupation?

100000000000000000000000000000000000000000	Architects / Surveyors
010000000000000000000000000000000000000000	Professional - Engineers
001000000000000000000000000000000000000000	Professional - Math and computer scientists
000100000000000000000000000000000000000000	Professional - Natural scientists
000010000000000000000000000000000000000000	Professional - Registered nurses
000001000000000000000000000000000000000000	Professional - Health diagnostics
000000100000000000000000000000000000000000	Professional - Other Health related
000000010000000000000000000000000000000000	Management - Accountants and auditors
000000001000000000000000000000000000000000	Management - Public sector executives & other managers
000000000100000000000000000000000000000000	Management - Private sector executives & other managers
000000000010000000000000000000000000000000	Management - All others
000000000001000000000000000000000000000000	Teachers
000000000000100000000000000000000000000000	Other professionals
000000000000010000000000000000000000000000	Technicians/technical - Engineering
000000000000001000000000000000000000000000	Technicians/technical- Health technologists & technicians
000000000000000100000000000000000000000000	Technicians/technical- Science technicians
000000000000000010000000000000000000000000	Technicians/technical - All other technicians
000000000000000001000000000000000000000000	Sales and related - Sales representatives
000000000000000000100000000000000000000000	Sales and related - Supervisors and proprietors
000000000000000000010000000000000000000000	Other sales related
000000000000000000001000000000000000000000	Adjustors and investors
000000000000000000000100000000000000000000	Administrative support - Computer operators
000000000000000000000010000000000000000000	Administrative support - Information clerks
000000000000000000000001000000000000000000	Administrative support - Secretaries
000000000000000000000000100000000000000000	Stenographers & typists
000000000000000000000000010000000000000000	Administrative support - Supervisors
000000000000000000000000000000000000000000	Other administrative support
000000000000000000000000001000000000000000	Craft and precision-production - Construction
000000000000000000000000000100000000000000	Craft and precision-production - All others
000000000000000000000000000010000000000000	Assemblers, operators, fabricators - Transportation
000000000000000000000000000001000000000000	Assemblers, operators, fabricators - Machine
000000000000000000000000000000100000000000	Assemblers, operators, fabricators-All others
000000000000000000000000000000010000000000	Cleaning equipment, handlers, laborers
000000000000000000000000000000001000000000	Service-related occupations - Personal
000000000000000000000000000000000100000000	Service-related occupations - Public safety
000000000000000000000000000000000010000000	Service-related occupations - Health
000000000000000000000000000000000001000000	Service-related occupations - Other
000000000000000000000000000000000000100000	Farmers, fisherman, foresters
000000000000000000000000000000000000010000	Other farming, fishing, & hunting
000000000000000000000000000000000000001000	Military
000000000000000000000000000000000000000100	Missing
000000000000000000000000000000000000000010	Not worked in the past 3 years
000000000000000000000000000000000000000001	Never worked

426

	Item #	Question	Contrast Coding	Variable Coding
SECTION E: Literacy Activities and Collaboration				
	E-1	How often do you read a newspaper?	00000 10000 01000 00100 00010 00001	Every day A few times a week Once a week Less than once a week Never Missing
Derived	E-1/ E-2	Reading Practice #1 Reads English newspaper at least once a week; sections: news, editorial, or financial	00 10 01	Yes No Missing
Derived		Reading Practice #2 Reads English newspaper at least once a week; sections: sports	00 10 01	Yes No Missing
Derived		Reading Practice #3 Reads English newspaper at least once a week; sections: home, fashion, health, reviews of movies, books and art	00 10 01	Yes No Missing
Derived		Reading Practice #4 Reads English newspaper at least once a week; sections: TV, movie, and concert listings; classified and other ads	00 10 01	Yes No Missing
Derived		Reading Practice #5 Reads English newspaper at least once a week; sections: comics, horoscopes, advice	00 10 01	Yes No Missing
	E-3	Number of different English magazines read on a regular basis.	10000 01000 00100 00000 00010 00001	None 1 2 3 - 5 6 or more Missing
	E-4	Different types of English books read recently. Fiction?	10 00 01	Yes No Missing
		Recreation or entertainment?	10 00 01	Yes No Missing
		Current affairs or history?	10 00 01	Yes No Missing
		Inspiration or religion?	10 00 01	Yes No Missing
		Science or social science?	10 00 01	Yes No Missing
		Reference, such as encyclopedias or dictionaries?	10 00 01	Yes No Missing
		Manuals for cooking, operating, repairing, or building?	10 00 01	Yes No Missing
		Any other type of books?	10 00 01	Yes No Missing
		None	00 10 01	No Yes Missing

	Item #	Question	Contrast Coding	Variable Coding
	E-5	Kinds of things read in English for own use. Letters or memos?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Reports, articles, magazines or journals?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Manuals or reference books?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Directions or instructions for medicines or recipes?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Diagrams or schematics?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Bills, invoices, spread-sheets, or budget tables?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	E-6	Kinds of things read in English as part of current/recent job. Letters or memos?	00010	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00000	Every day
			00001	Missing
		Reports, articles, magazines or journals?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Manuals or reference books?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Directions or instructions for medicines or recipes?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing

	Item #	Question	Contrast Coding	Variable Coding
	E-6 (contd.)	Diagrams or schmatics?	10000	Never
			00000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Bills, invoices, spread-sheets, or budget tables?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	E-7	Things that you write or fill out in English for own use	10000	Never
			00000	Less than once a week
		Letters or memos?	01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Forms such as bills, invoices, or budgets?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Reports or articles?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	E-8	Things that you write or fill out in English as part of job	00010	Never
			10000	Less than once a week
		Letters or memos?	01000	Once a week
			00100	A few times a week
			00000	Every day
			00001	Missing
		Forms such as bills, invoices, or budgets?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Reports or articles?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	E-9	How often do you use arithmetic or math for your own use?	00010	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00000	Every day
			00001	Missing
	E-10	How often do you use arithmetic or math as part of job?	00010	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00000	Every day
			00001	Missing
	E-11	How much help do you get from family or friends with filling out forms?	0000	None
			1000	A little
			0100	Some
			0010	A lot
			0001	Missing

	Item #	Question	Contrast Coding	Variable Coding
	E-11 (contd.)	reading or explaining newspaper articles/written information?	0000 1000 0100 0010 0001	None A little Some A lot Missing
		printed information from government agencies, public companies, etc?	0000 1000 0100 0010 0001	None A little Some A lot Missing
		writing notes or letters?	0000 1000 0100 0010 0001	None A little Some A lot Missing
		using basic arithmetic, as in filling out order forms, or balancing a checkbook?	0000 1000 0100 0010 0001	None A little Some A lot Missing

Appendix B: Conditioning Variables
NATIONAL ADULT LITERACY SURVEY - PRISON SAMPLE
Contrast Coding for the MGROUP Conditioning Variables

	Item #	Question	Contrast Coding	Variable Coding
	B-1	Highest level of education completed in the USA?	100000000000	Less than high school
			010000000000	Some high school
			001000000000	GED or high school equivalency
			000000000000	High school graduate
			000100000000	Vocational
			000010000000	College: less than 2 years
			000001000000	College: Associate's degree
			000000100000	College: two years or more no degree
			000000010000	College graduate (B.S. or B.A.)
			000000001000	Postgraduate, no degree
			000000000100	Postgraduate degree
			000000000010	Don't know
			000000000001	Missing
	B-3	Studied for a GED or high school equivalency?	00	Yes
			10	No
			01	Missing
SECTION F: Demographic Information				
	F-1	When you were growing up, who did you live with most of the time?	100000000	Mother
			010000000	Father
			000000000	Both parents
			001000000	Grandparents
			000100000	Other relatives
			000010000	Friends
			000001000	Foster homes
			000000100	Agency or institution
			000000010	Someone else
			000000001	Missing
	F-2	Highest level of education your mother completed?	100000000000	Less than high school
			010000000000	Some high school
			001000000000	GED or high school equivalency
			000000000000	High school graduate
			000100000000	Vocational
			000010000000	College: less than 2 years
			000001000000	College: Associate's degree
			000000100000	College: two years or more no degree
			000000010000	College graduate (B.S. or B.A.)
			000000001000	Postgraduate, no degree
			000000000100	Postgraduate degree
			000000000010	Don't know
			000000000001	Missing
	F-3	Highest level of education your father completed?	100000000000	Less than high school
			010000000000	Some high school
			001000000000	GED or high school equivalency
			000000000000	High school graduate
			000100000000	Vocational
			000010000000	College: less than 2 years
			000001000000	College: Associate's degree
			000000100000	College: two years or more no degree
			000000010000	College graduate (B.S. or B.A.)
			000000001000	Postgraduate, no degree
			000000000100	Postgraduate degree
			000000000010	Don't know
			000000000001	Missing
	F-4	Current marital status?	0000	Never married
			1000	Married
			0100	Separated or divorced
			0010	Widowed
			0001	Missing

	Item #	Question	Contrast Coding	Variable Coding
Derived	F-5	Race/Ethnicity?	0000000000	Black not Hispanic
			1000000000	Hispanic, Mexicano, Mexican, Mexican American, or Chicano
			0100000000	Hispanic, Puerto Rican
			0010000000	Hispanic, Cuban
			0001000000	Hispanic, Central or South American or other Spanish
			0000100000	Asian
			0000010000	American Indian or Alaskan Native
			0000001000	Pacific Islander
			0000000100	White
			0000000010	Other
			0000000001	Missing
	F-6	Are you of Spanish or Hispanic origin or descent?	10	Yes
			00	No
			01	Missing
	F-7	Which best describes your Hispanic origin?	00000	Mexicano, Mexican, Mexican American, or Chicano
			10000	Puerto Rican
			01000	Cuban
			00100	Central or South American
			00010	Other Spanish or Hispanic
			00001	Missing
	F-8	Date of birth		Actual age Range 16 – 64
	F-9	Sex	00	Male
			10	Female
			01	Missing
		Census Region of the country	100	Northeast
			000	Midwest
			010	South
			001	West
SECTION C: Current Offenses and Criminal History				
	C-1	C-1 Record of offenses--not scored		
	C-2	For which of these offenses did you receive the longest sentence?	100000000000	Murder, code 11
			010000000000	Rape including attempted rape, code 12
			000000000000	Armed robbery, code 13
			001000000000	Assault and other violent crimes, codes 14 & 15
			000100000000	Burglary, code 21
			000010000000	Auto theft and other property offenses, codes 22 & 26
			000001000000	Forgery, fraud, and embezzlement, code 23
			000000100000	Larceny, grand, and petty, code 24
			000000010000	Stolen property, code 25
			000000001000	Drugs, code 31
			000000000100	Public order, probation violation, rioting, etc., code 41
			000000000010	Federal offenses, code 51
			000000000001	Missing, 88 Don't know, and 99 omit.
Derived	C-3	In what month and year were you first arrested?	Time in years	Administration date (9/1/92) minus arrest date, rounded to years Missing = mean Don't know = mean
Derived	C-4	In what month and year were you first admitted to prison?	Time in years	Administration date (9/1/92) minus admission date, rounded to years Missing = mean Don't know = mean

	Item #	Question	Contrast Coding	Variable Coding
	C-5	Have you ever been placed on probation, either as a juvenile or as an adult?	10	No
			00	Yes
			01	Missing
Derived	C-6	How many times have you ever been placed on probation as a juvenile?	100000	None
			000000	1 time
			010000	2 times
			001000	3 - 5 times
			000100	6+ times
			000010	Don't know
			000001	Missing
Derived	C-7	How many times have you been placed on probation as an adult?	100000	None
			000000	1 time
			010000	2 times
			001000	3 - 5 times
			000100	6+ times
			000010	Don't know
			000001	Missing
	C-8	Did you ever serve time in prison, jail, or some other correctional facility for any other offenses?	10	No
			00	Yes
			01	Missing
Derived	C-9	How many times as a juvenile?	100000	None
			000000	1 time
			010000	2 times
			001000	3 - 5 times
			000100	6+ times
			000010	Don't know
			000001	Missing
Derived	C-10	How many times as an adult?	100000	None
			000000	1 time
			010000	2 times
			001000	3 - 5 times
			000100	6+ times
			000010	Don't know
			000001	Missing
	C-11	Do you have a definite date on which you expect to be released?	10	Yes
			00	No
			01	Missing
Derived	C-12	In what month and year will you be released?	Time in years	Release date minus administration date (9/1/92), rounded to years Missing = mean Don't know = mean
	C-13	Do you expect to ever be released from prison?	00	Yes
			10	No
			01	Missing
Derived	C-14	In what month and year is your earliest possible release date?	Time in years	Earliest possible release date minus administration date, rounded to years Missing = mean Don't know = mean
SECTION A: General and Language Background				
Derived	A-1	In what country were you born?	00000	USA
			10000	Spanish speaking countries
			01000	European language countries
			00100	Asian language countries
			00010	All other countries
			00001	Missing

	Item #	Question	Contrast Coding	Variable Coding
	A-2	How many years have you lived in the United States?	1	1 - 5 years
			2	6 - 10 years
			3	11 - 15 years
			4	16 - 20 years
			5	21 - 30 years
			6	31 - 40 years
			7	41 - 50 years
			8	51 or more years
			9	Missing by design - skip pattern
Derived	A-2b	Age of arrival in the USA (Age minus median of A-2)	0	Born in the USA or Missing
			0	Negative values
				Age 1 - 90
	A-3	Highest education before coming to the USA?	100	Born in the USA or missing by design
			000	Primary or elementary (codes 2 & 3)
			010	Some schooling (codes 4 to 7)
			001	No schooling (code 1)
Derived	A-4	Languages spoken in the home while growing up?	000000000	English only
			100000000	English & Spanish
			010000000	English & European
			001000000	English & Asian
			000100000	Spanish & Other
			000010000	Other & Other
			000001000	Spanish only
			000000100	European only
			000000010	Asian only
			000000001	Missing
Derived	A-5	Languages learned to speak before starting school?	000000000	English only
			100000000	English & Spanish
			010000000	English & European
			001000000	English & Asian
			000100000	Spanish & Other
			000010000	Other & Other
			000001000	Spanish only
			000000100	European only
			000000010	Asian only
			000000001	Missing
Derived	A-6	What language did you first learn to read and write?	000000	English
			100000	Spanish
			010000	European
			001000	Asian
			000100	Other Non-English language
			000010	Never learned to read and write
			000001	Missing
	A-7	How old were you when you learned to speak English?	000000	1 - 4 years old
			100000	5 - 10 years old
			010000	11 - 15 years old
			001000	16 - 20 years old
			000100	21 or older
			000010	Does not speak English
			000001	Missing
	A-8	With regard to (NON-ENGLISH LANGUAGE IN A-5) how well do you understand it when it is spoken to you?	1000	Not at all
			0100	Not well
			0000	Well
			0010	Very well
			0001	Missing
		speak it?	1000	Not at all
			0100	Not well
			0000	Well
			0010	Very well
			0001	Missing

	Item #	Question	Contrast Coding	Variable Coding
	A-8 (contd.)	read it?	1000 0100 0000 0010 0001	Not at all Not well Well Very well Missing
		write it?	1000 0100 0000 0010 0001	Not at all Not well Well Very well Missing
	A-9	With regard to (NON-ENGLISH LANGUAGE IN A-5) how often do you listen to radio programs in that language?	10000 00000 01000 00100 00010 00001	Never Less than once a week Once a week A few times a week Every day Missing
		watch television programs in that language?	10000 00000 01000 00100 00010 00001	Never Less than once a week Once a week A few times a week Every day Missing
		read newspapers, magazines or books in that language?	10000 00000 01000 00100 00010 00001	Never Less than once a week Once a week A few times a week Every day Missing
		write or fill out letters or forms in that language?	10000 00000 01000 00100 00010 00001	Never Less than once a week Once a week A few times a week Every day Missing
	A-10	Have you taken a course to learn how to read and write English (2nd lang)?	00 10 01	Yes No Missing
	A-11	Did you complete this course?	00 10 01	Yes No Missing
	A-12	Have you taken a course to learn to speak or understand English (2nd lang)?	00 10 01	Yes No Missing
	A-13	Did you complete this course?	00 10 01	Yes No Missing
Derived	A-14	Which language do you usually speak now?	00000 10000 01000 00100 00010 00001	English Spanish European Asian Other Non-English language Missing
Derived	A-15	What other language do you often speak now?	000000 100000 010000 001000 000100 000010 000001	English Spanish European Asian Other Non-English language No other language spoken Missing

435

	Item #	Question	Contrast Coding	Variable Coding
	A-16	With regard to the English language, how well do you understand it when it is spoken to you?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		speak it?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		read it?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		write it?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
		do arithmetic problems?	1000	Not at all
			0100	Not well
			0010	Well
			0000	Very well
			0001	Missing
SECTION B: Educational Background and Experiences				
	B-2	Main reason you stopped your schooling?	1000000	Financial problems
			0100000	Went to work or into the military
			0010000	Pregnancy
			0000000	Lost interest or behavior problems, and academic problems
			0001000	Family or personal problems
			0000100	Convicted of crime or sent to jail, prison, or detention center
			0000010	Other
			0000001	Missing
	B-4	Kind of high school program?	1000	General
			0000	Vocational, technical, or trade
			0100	College preparatory
			0010	The same for everyone
			0001	Missing
	B-5	Since your current admission to prison, have you ever been in any education program, excluding vocational training?	00	Yes
			10	No
			01	Missing
	B-6	What kind of program was that? Basic classes up to 9th grade	00	No
			10	Yes
			01	Missing
		High school classes for diploma or GED	00	No
			10	Yes
			01	Missing
		College-level classes	00	No
			10	Yes
			01	Missing
		Other	00	No
			10	Yes
			01	Missing
	B-7	Are you still in an education program?	00	Yes
			10	No
			01	Missing

	Item #	Question	Contrast Coding	Variable Coding
	B-8	Since your current admission to prison, how long have you spent altogether in education program(s)?	1000 0000 0100 0010 0001	Number of years Number of months Less than one month Don't know Missing
Derived	B-8	Specified time spent in educational program	0 Number of months	Missing & Don't know Converted years to months
	B-9	Since your current admission to prison, have you ever been in any vocational training program, excluding prison work assignments?	0 10 01	Yes No Missing
	B-10	Are you still in a vocational program?	00 10 01	Yes No Missing
	B-11	Since your current admission to prison, how long have you spent altogether in vocational training?	1000 0000 0100 0010 0001	Number of years Number of months Less than one month Don't know Missing
Derived	B-11	Specified time spent in vocational program	0 Number of months	Missing and Don't know Converted years to months
	B-12	Since your current admission to prison, did you ever take part in a program other than in regular school in order to improve your basic skills, that is, basic reading, writing, and arithmetic skills?	00 10 01	Yes No Missing
	B-13	Was this program training program or course given or sponsored by your employer or union?	000 100 010 001	Yes No Don't know Missing
		publicly sponsored training program such as JTPA or ABE?	000 100 010 001	Yes No Don't know Missing
		tutoring program sponsored by library, church, or community?	000 100 010 001	Yes No Don't know Missing
		any other program?	000 100 010 001	Yes No Don't know Missing
	B-14	If 'Yes' to B-13, when taken? employer training program	000 100 010 001	Within the past year 1 - 5 years ago More than five years ago Missing
		publicly sponsored program, JTPA or ABE	000 100 010 001	Within the past year 1 - 5 years ago More than five years ago Missing
		tutoring program by library, church or community	000 100 010 001	Within the past year 1 - 5 years ago More than five years ago Missing
		any other program, military, prisons or institutions	000 100 010 001	Within the past year 1 - 5 years ago More than five years ago Missing

	Item #	Question	Contrast Coding	Variable Coding
	B-15	Do you have a physical, mental, or other health condition that keeps you from participating fully in work, school, or other activities?	00 10 01	No Yes Missing
	B-16	Do you have any difficulty seeing the words and letters in ordinary newspaper print even when wearing glasses or contact lenses, if you usually wear them?	00 10 01	No Yes Missing
	B-17	Difficulty hearing a normal conversation even when using a hearing aid, if you usually wear one?	00 10 01	No Yes Missing
	B-18	Do you currently have any of these conditions? A learning disability?	00 10 01	No Yes Missing
		Any mental or emotional condition?	00 10 01	No Yes Missing
		Mental retardation?	00 10 01	No Yes Missing
		A speech disability?	00 10 01	No Yes Missing
		A physical disability?	00 10 01	No Yes Missing
		A long-term illness (6 months or more)?	00 10 01	No Yes Missing
		Any other health impairment?	00 10 01	No Yes Missing
SECTION D: Prison Work Assignments and Labor Force Participation				
	D-1	Do you currently have any work assignments inside or outside prison?	00 10 01	Yes No Missing
	D-2	What work assignments do you have now? Goods production?	10 00 01	Yes No Missing
		General janitorial?	10 00 01	Yes No Missing
		Grounds or road maintenance?	10 00 01	Yes No Missing
		Food preparation or related duties	10 00 01	Yes No Missing
		Laundry?	10 00 01	Yes No Missing
		Hospital, infirmary, or other medical services?	10 00 01	Yes No Missing

	Item #	Question	Contrast Coding	Variable Coding
	D-2 (contd.)	Farming, forestry, or ranching?	10	Yes
			00	No
			01	Missing
		Other services (library, stockroom, store, etc.)?	10	Yes
			00	No
			01	Missing
		Maintenance, repair, or construction?	10	Yes
			00	No
			01	Missing
		Enrolled in school?	10	Yes
			00	No
			01	Missing
		Other?	10	Yes
			00	No
			01	Missing
Derived	D-3	In the last week how many hours were you assigned to these jobs?	10000	0 - 19 hours
			01000	20 - 29 hours
			00100	30 - 39 hours
			00000	40 hours
			00010	41 + hours
			00001	Missing and 999
	D-4	Are you paid money for any of this work?	00	Yes
			10	No
			01	Missing
Derived	D-5	How much did you make last month?	00000	$1 - $19
			10000	$20 - $29
			01000	$30 - $39
			00100	$40 - $59
			00010	$60 +
			00001	Missing and 99999
	D-6	Other than money, do you receive anything for work, such as time credits or other privileges?	00	Yes
			10	No
			01	Missing
	D-7	What do you receive? Good time?	00	Yes
			10	No
			01	Missing
		Cigarettes, food, etc.?	10	Yes
			00	No
			01	Missing
		Extra privileges?	10	Yes
			00	No
			01	Missing
		Other?	10	Yes
			00	No
			01	Missing
	D-8	In the year before your arrest did you receive income from any of the following? Social Security or other pension?	10	Yes
			00	No
			01	Missing
		Welfare or charity?	10	Yes
			00	No
			01	Missing
		Unemployment or worker's compensation?	10	Yes
			00	No
			01	Missing
		Family or friends?	10	Yes
			00	No
			01	Missing

	Item #	Question	Contrast Coding	Variable Coding
	D-8 (contd.)	Illegal sources?	10	Yes
			00	No
			01	Missing
		Educational scholarship or grant?	10	Yes
			00	No
			01	Missing
		Something else?	10	Yes
			00	No
			01	Missing
		None?	10	Yes
			00	No
			01	Missing
	D-9	In the past three years, did you work for pay or profit, while not serving time in prison?	00	Yes
			10	No
			01	Missing
Derived	D-10	Between January 1989 and your admission to prison, how many months did you work?	1 month	Less than 1
			0	Missing
				Actual number of months
Derived	D-11	What was your business or industry prior to your current admission to prison?	10000000	Agriculture, forestry, fishing code 1
			01000000	Construction & Mining codes 2, & 3
			00100000	Manufacturing - Durable or nondurable codes 4, & 5
			00010000	Transportation, Communications, Utilities codes 6, 7, & 8
			00001000	Wholesale or retail trade codes 9, 10, 11, & 12
			00000100	Finance - Banking & Real estate codes 13, & 15
			00000010	Services-Bus., Pers., Entertain., Profess. codes 16, 17, 18, & 19
			00000001	Public Administration & Military codes 20, & 21
			00000000	Missing or Not worked in past 3 years codes 22, & 23
Derived	D-12	What was your occupation prior to your current admission to prison?	1000000000000	Other professional codes 2, 10, 11, 1313
			0100000000000	Sales and related codes 18, 19, 2020
			0010000000000	Administrative support & Other technicians codes 14, 15, 17, 21, 22, 24, 26, 2727
			0001000000000	Craft and precision-production - Construction crafts code 28
			0000100000000	Craft and precision-production - All others code 29
			0000010000000	Assemblers, operators, fabricators - Transport oper. code 30
			0000001000000	Assemblers, operators, fabricators - Fabricators code 31
			0000000100000	Assemblers, operators, fabricators - All others code 32
			0000000010000	Cleaning equipment, handlers, laborers code 33
			0000000001000	Service-related occupations - Personal code 34
			0000000000100	Service-related occupations – Other codes 35, 36, 37, 4037
			0000000000010	Managers & Operators codes 38, 3939
			0000000000001	Missing code 41
			0000000000000	Not worked in past 3 years code 42

	Item #	Question	Contrast Coding	Variable Coding
	D-14	Was this full-time, part-time, or occasional work?	000 100 010 001	Full-time Part-time Occasional Missing
	D-15	What were your average monthly earnings before any deductions?	10000000000000 01000000000000 00100000000000 00010000000000 00001000000000 00000100000000 00000010000000 00000000000000 00000001000000 00000000100000 00000000010000 00000000001000 00000000000100 00000000000010 00000000000001	$1 - $99 $100 - $199 $200 - $299 $300 - $399 $400 - $499 $500 - $599 $600 - $799 $800 - $999 $1,000 - $1,199 $1,200 - $1,499 $1,500 - $1,999 $2,000 - $2,499 $2,500 - $4,999 $5,000 or more Don't Know, No Income, and Missing
SECTION E: Literacy Activities and Collaboration				
	E-1	How often do you read a newspaper?	00000 10000 01000 00100 00010 00001	Every day A few times a week Once a week Less than once a week Never Missing
Derived	E-1/ E-2	Reading Practice #1 Reads English newspaper at least once a week; sections: news, editorial, or financial	00 10 01	Yes No Missing
Derived		Reading Practice #2 Reads English newspaper at least once a week; sections: sports	00 10 01	Yes No Missing
Derived		Reading Practice #3 Reads English newspaper at least once a week; sections: home, fashion, health, reviews of movies, books and art	0 10 01	Yes No Missing
Derived		Reading Practice #4 Reads English newspaper at least once a week; sections: TV, movie, and concert listings; classified and other ads	00 10 01	Yes No Missing
Derived		Reading Practice #5 Reads English newspaper at least once a week; sections: comic, horoscopes, advice	00 10 01	Yes No Missing
	E-3	Number of different English magazines read on a regular basis.	10000 01000 00100 00000 00010 00001	None 1 2 3 - 5 6 or more Missing
	E-4	Different types of English books read recently. Fiction?	10 00 01	Yes No Missing
		Recreation or entertainment?	10 00 01	Yes No Missing

	Item #	Question	Contrast Coding	Variable Coding
	E-4 (contd.)	Current affairs or history?	10	Yes
			00	No
			01	Missing
		Inspiration or religion	10	Yes
			00	No
			01	Missing
		Science or social science?	10	Yes
			00	No
			01	Missing
		Reference, such as encyclopedias or dictionaries?	10	Yes
			00	No
			01	Missing
		Manuals for cooking, operating, repairing, or building?	10	Yes
			00	No
			01	Missing
		Any other type of books?	10	Yes
			00	No
			01	Missing
		None	00	No
			10	Yes
			01	Missing
	E-5	Kinds of things read in English for own use. Letters or memos?	10000	Never
			01000	Less than once a week
			00100	Once a week
			00000	A few times a week
			00010	Every day
			00001	Missing
		Reports, articles, magazines or journals?	10000	Never
			01000	Less than once a week
			00100	Once a week
			00000	A few times a week
			00010	Every day
			00001	Missing
		Manuals or reference books?	10000	Never
			01000	Less than once a week
			00100	Once a week
			00000	A few times a week
			00010	Every day
			00001	Missing
		Directions or instructions for medicines or recipes?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Diagrams or schematics?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Bills, invoices, spread-sheets, or budget tables?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	E-6	Things that you write or fill out in English for own use. Letters or memos?	01000	Never
			00100	Less than once a week
			00000	Once a week
			00010	A few times a week
			00001	Every day
			10000	Missing

	Item #	Question	Contrast Coding	Variable Coding
	E-6 (contd.)	Forms?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
		Reports or articles?	00000	Never
			10000	Less than once a week
			01000	Once a week
			00100	A few times a week
			00010	Every day
			00001	Missing
	E-7	How often do you use arithmetic or mathematics?	10000	Never
			01000	Less than once a week
			00100	Once a week
			00010	A few times a week
			00000	Every day
			00001	Missing
	E-8	How much help do you get from family or friends with filling out forms?	0000	None
			1000	A little
			0100	Some
			0010	A lot
			0001	Missing
		reading or explaining newspaper articles or written information?	0000	None
			1000	A little
			0100	Some
			0010	A lot
			0001	Missing
		printed information from government agencies, public companies, etc.?	0000	None
			1000	A little
			0100	Some
			0010	A lot
			0001	Missing
		writing notes or letters?	0000	None
			1000	A little
			0100	Some
			0010	A lot
			0001	Missing
		using basic arithmetic, as in filling out order forms, or balancing a checkbook?	0000	None
			1000	A little
			0100	Some
			0010	A lot
			0001	Missing
	E-9	Since your admission to prison have you joined any organization authorized by prison authorities?	00	Yes
			10	No
			01	Missing
	E-10	What groups have you joined? Drug awareness or other drug dependency group?	10	Yes
			00	No
			01	Missing
		Alcoholics anonymous?	10	Yes
			00	No
			01	Missing
		Bible club or other religious group?	10	Yes
			00	No
			01	Missing
		Ethnic/racial organization?	10	Yes
			00	No
			01	Missing
		Prisoner assistance group?	10	Yes
			00	No
			01	Missing

	Item #	Question	Contrast Coding	Variable Coding
	E-10 (contd.)	Other prisoner self-help or personal improvement group?	10 00 01	Yes No Missing
		Outside community activities?	10 00 01	Yes No Missing
		Religious activities?	10 00 01	Yes No Missing
		Classes in parenting and childrearing skills?	10 00 01	Yes No Missing
		Classes in life skills?	10 00 01	Yes No Missing
		Classes in art or crafts?	10 00 01	Yes No Missing
		Prerelease programs?	10 00 01	Yes No Missing
		Other?	10 00 01	Yes No Missing

Appendix C

GAMMAS

Appendix C: GAMMAs
NATIONAL ADULT LITERACY SURVEY
Main and Non-Incentive Household Samples, Age 16–64

	Prose	Document	Quantitative		Prose	Document	Quantitative
1	0.265335	0.804745	0.105941	57	-0.016558	-0.015372	-0.023200
2	0.089618	0.077778	0.078741	58	0.003645	0.019359	0.015859
3	-0.098994	-0.099482	-0.108083	59	-0.026824	-0.029882	-0.027388
4	-0.085802	-0.073334	-0.075899	60	0.024087	0.020528	0.012898
5	0.005281	0.009186	0.008700	61	0.005785	0.011939	0.019368
6	0.058752	0.052258	0.058943	62	-0.009555	0.000500	-0.001404
7	-0.000658	-0.012959	0.013640	63	-0.027453	-0.040464	-0.042397
8	-0.081927	-0.078919	-0.081139	64	-0.016195	-0.018048	-0.013587
9	-0.062319	-0.050080	-0.027659	65	0.013867	0.008309	0.009102
10	-0.005154	0.005567	0.004550	66	-0.011886	-0.014432	0.000339
11	0.032474	0.018837	0.025130	67	0.015101	0.015852	0.000861
12	0.087153	0.063777	0.063273	68	-0.009220	-0.008507	0.008988
13	-0.041089	-0.043368	-0.058670	69	0.015646	0.019675	0.037311
14	0.010729	0.010027	0.010322	70	0.025950	0.027008	0.015920
15	0.008652	0.012522	0.027859	71	0.007439	0.006329	-0.003490
16	0.011704	0.010370	-0.000743	72	-0.001630	-0.007678	-0.016148
17	-0.043997	-0.036312	-0.024674	73	0.002593	0.000001	0.009350
18	-0.008761	-0.000211	-0.000674	74	0.020148	0.026341	0.031159
19	0.029202	0.018288	0.059150	75	0.028416	0.032969	0.035664
20	0.011303	0.020183	0.017691	76	-0.009798	-0.011303	-0.016582
21	0.034129	0.047237	0.036427	77	0.009450	0.005016	0.015584
22	-0.006446	0.019028	0.016070	78	0.003760	0.004124	0.017895
23	0.049728	0.064333	0.053318	79	0.001562	-0.003232	0.009698
24	0.014517	0.022358	0.025017	80	-0.000347	0.008695	0.006219
25	-0.033761	-0.024672	-0.021914	81	0.005261	0.012445	0.007496
26	0.009857	-0.001147	-0.003722	82	-0.005057	-0.002025	-0.005655
27	0.017347	0.014354	0.001806	83	-0.006922	-0.012512	-0.010622
28	-0.041717	-0.037384	-0.042831	84	-0.001988	-0.000527	-0.004965
29	-0.018460	-0.019807	-0.004530	85	0.021699	0.010949	-0.004407
30	-0.028313	-0.015228	-0.021845	86	0.001572	0.009925	0.021979
31	0.022586	0.019478	0.015017	87	-0.000828	-0.003077	-0.008783
32	0.003435	0.015407	0.022653	88	0.013358	0.002082	0.003669
33	-0.025806	-0.038468	-0.036369	89	-0.002125	0.001826	0.004763
34	0.039053	0.037766	0.039987	90	-0.002163	-0.004323	-0.002081
35	-0.050330	-0.055816	-0.038664	91	0.006828	0.003089	-0.000441
36	-0.003240	-0.012636	-0.023190	92	-0.015072	-0.012641	-0.022985
37	-0.028447	-0.032116	-0.039043	93	-0.008271	-0.003251	0.000932
38	-0.029767	-0.038622	-0.045725	94	-0.011197	-0.013226	-0.010896
39	-0.011398	-0.017870	-0.022668	95	0.010933	0.018050	0.012456
40	0.019381	0.031753	0.018726	96	-0.002232	0.001844	0.010344
41	-0.012588	-0.015410	-0.028513	97	0.004850	-0.005177	-0.005398
42	-0.018690	-0.007193	-0.003818	98	0.002923	0.001516	-0.009571
43	-0.004112	-0.006297	0.000534	99	-0.005280	0.012089	0.000964
44	-0.008020	-0.011730	-0.012945	100	-0.002985	0.002226	0.003561
45	-0.008310	-0.006569	-0.015524	101	0.007666	0.011213	0.021204
46	-0.006556	-0.001768	-0.005404	102	-0.003482	-0.002242	0.011794
47	-0.004023	-0.006307	-0.006962	103	0.000789	-0.000146	-0.006200
48	0.013940	0.012422	0.017807	104	-0.004100	-0.004462	-0.003605
49	0.033373	0.023934	0.023238	105	0.000138	-0.024545	-0.010357
50	-0.026636	-0.029764	-0.016504	106	0.014426	0.004010	0.010523
51	-0.001077	0.021235	0.009577	107	0.001232	-0.003193	0.004744
52	-0.010024	-0.013072	-0.004162	108	-0.014275	-0.013820	-0.016452
53	-0.003845	0.004828	0.016756	109	-0.011148	-0.008718	-0.013966
54	-0.037755	-0.033688	-0.039147	110	-0.009086	-0.020447	-0.025463
55	0.024290	0.015148	0.019197	111	-0.005455	0.007377	-0.000715
56	-0.007125	-0.000717	-0.008696	112	0.007569	-0.005231	-0.002140

	Prose	Document	Quantitative
113	-0.002373	0.010212	-0.001589
114	-0.016543	-0.021711	-0.011960
115	0.004638	-0.007402	-0.000868
116	0.011453	0.036551	0.025036
117	0.006720	0.010553	0.004190
118	0.009637	0.005190	-0.001368
119	0.010283	-0.003781	-0.009548
120	-0.017395	-0.010387	-0.016695
121	0.005257	-0.006465	0.008145
122	-0.020572	-0.002708	-0.001211
123	-0.012164	-0.012665	-0.021782
124	0.004722	0.001577	-0.004611
125	-0.002616	-0.010286	0.001642
126	0.004285	0.007716	0.005620
127	-0.016692	-0.014794	-0.030812
128	-0.002551	-0.012456	-0.005198
129	-0.022426	-0.019560	-0.014941
130	-0.009327	-0.004229	-0.004523
131	0.006057	-0.005847	-0.004935
132	-0.024649	-0.015746	-0.025266
133	0.029475	0.028590	0.022296
134	-0.002060	-0.005230	0.001587
135	0.026443	0.026166	0.015835
136	-0.004092	-0.011434	-0.003882
137	0.008297	0.004666	-0.001050
138	-0.017263	-0.004941	-0.011137
139	0.013212	0.000054	0.023319
140	0.000738	0.011920	-0.002400
141	-0.007818	-0.010457	-0.004738
142	-0.006186	0.007958	-0.004542
143	0.006922	0.000524	0.008679
144	0.019380	0.016977	0.016165
145	-0.016618	-0.032579	-0.036051
146	0.000468	-0.008613	-0.001458
147	-0.004572	-0.008013	-0.005281
148	0.001343	0.004474	0.017433
149	0.009789	0.005964	0.002163
150	0.022522	0.030207	0.020913
151	0.005286	-0.011300	-0.009045
152	-0.005086	0.001963	-0.004316
153	-0.004884	0.003073	-0.005555
154	0.016227	0.018133	0.010827
155	-0.006091	-0.001456	-0.010085
156	0.013045	0.010626	0.012998
157	-0.014362	0.004076	-0.003217
158	0.002333	0.000109	-0.008684
159	0.004217	-0.000414	-0.003287
160	0.009967	0.000658	0.014707
161	0.002900	-0.003366	-0.004038
162	0.003034	0.023169	0.010434
163	0.015120	0.007044	0.007719
164	0.028756	0.014615	0.022283
165	-0.015669	-0.019315	-0.026997
166	0.005442	0.002427	0.010632
167	0.011940	0.001765	0.007562
168	0.001273	0.009385	0.007002
169	0.002047	0.012252	0.015858
170	-0.005896	-0.024218	-0.016733
171	-0.005283	-0.022468	-0.028083
172	0.019783	0.015767	0.020910
173	-0.004697	-0.009615	-0.018191
174	0.015362	0.024345	0.015365
175	-0.021007	-0.020012	-0.020810
176	-0.023900	-0.022856	-0.020649
177	-0.002392	-0.005658	-0.017036
178	0.008998	0.020790	0.007610
179	-0.006965	0.003941	-0.000717
180	0.001260	-0.001553	-0.002255
181	0.024210	0.014705	0.023875
182	0.008997	0.008810	0.012061
183	-0.010504	-0.012956	-0.012596
184	-0.001276	0.001444	0.006480
185	-0.000642	-0.000385	-0.003270
186	-0.008008	-0.007024	-0.010142
187	0.005563	0.007239	-0.001017
188	0.000251	0.006956	-0.010031
189	-0.011760	-0.000416	-0.002364
190	-0.002156	-0.009894	-0.004488
191	-0.008394	-0.008402	-0.022475
192	0.006535	0.002823	0.011733
193	-0.003157	-0.001385	-0.001303
194	0.020533	0.020609	0.019994
195	0.004248	0.001893	-0.021445
196	-0.009171	-0.004867	-0.010794
197	0.015261	0.010885	0.026922
198	0.008910	0.005376	0.016380
199	0.010186	-0.009126	-0.008352
200	0.011078	0.015655	0.010602
201	0.020261	0.011497	0.002400
202	-0.003115	-0.003894	0.005149
203	0.013163	-0.002083	-0.020769
204	0.003308	0.001930	0.011198
205	-0.011342	-0.008792	-0.004710
206	0.002380	0.013885	0.008509
207	0.018905	0.007388	0.009025
208	0.001057	-0.009524	-0.019097
209	0.004091	0.014579	0.023635
210	-0.014726	-0.006385	0.006982
211	-0.000720	-0.006525	0.000441
212	0.016831	0.013602	0.013155
213	-0.007249	-0.006123	-0.011611
214	-0.001835	0.010296	0.026188
215	0.017526	0.014372	0.021485
216	0.027479	0.015801	0.006711
217	-0.003637	0.005876	0.001814
218	-0.002792	-0.011310	-0.008270
219	0.007224	0.023801	0.010866
220	0.020661	-0.001972	0.019749
221	-0.016737	-0.017831	-0.005839
222	0.003731	-0.020808	-0.016815
223	-0.013608	-0.016796	0.004252
224	0.010039	-0.002524	0.008295
225	-0.002707	0.000931	0.008868
226	-0.004389	-0.006907	-0.016753
227	-0.014123	-0.007810	-0.008849
228	-0.011249	-0.000428	-0.004609
229	0.009238	-0.001673	0.000976
230	0.002676	0.002294	-0.003061
231	0.009788	0.014611	0.004988
232	0.010464	0.009747	0.018519
233	-0.016491	-0.018010	-0.031394
234	0.003547	-0.005519	-0.004862
235	0.000365	-0.001456	-0.002909
236	0.024813	0.027695	0.007044
237	-0.017972	-0.000101	-0.011414
238	-0.001945	-0.000868	0.016459

	Prose	Document	Quantitative
239	-0.004129	-0.001717	-0.007502
240	-0.002757	0.006748	0.005374
241	0.003716	0.000664	0.013437
242	0.007634	0.016790	0.005115
243	0.008672	0.020240	0.011483
244	0.006144	0.009744	0.024341
245	0.002872	0.002807	-0.001361
246	-0.001543	0.013205	0.010628
247	0.003938	0.012877	0.027139
248	0.017789	0.025088	0.032337
249	0.027428	0.033702	0.016785
250	0.018357	0.012784	0.007766
251	0.003549	0.009985	-0.005304
252	0.000893	-0.004366	0.015152
253	0.007625	0.013527	0.002629
254	-0.007491	-0.008625	-0.019768
255	0.019562	0.007007	0.009167
256	-0.009684	-0.020263	-0.015248
257	-0.017160	-0.016521	-0.015739
258	0.007790	0.015418	0.016446
259	-0.009416	0.003036	0.010791
260	-0.013065	0.003506	-0.009276
261	-0.004983	-0.009331	-0.007449
262	-0.001545	0.003075	0.008132
263	-0.021977	-0.025129	-0.028226
264	-0.020404	-0.015143	0.000040
265	-0.019352	0.000277	0.007234
266	0.013021	-0.004487	0.005039
267	-0.023849	-0.023053	-0.006533
268	0.004730	0.001252	-0.000934
269	-0.015473	-0.001472	-0.011120
270	-0.009401	-0.002765	0.000947
271	0.005993	-0.006015	-0.000862
272	-0.015672	-0.007736	0.006670
273	-0.006970	-0.006553	-0.001956
274	-0.002355	0.002409	-0.002667
275	0.012721	-0.006334	0.013836
276	-0.001624	-0.002777	-0.017361
277	0.001370	0.001784	0.006778
278	-0.023137	-0.035461	-0.030340
279	-0.003975	-0.015062	-0.010293
280	0.011946	-0.009376	0.005789
281	0.005134	0.006491	-0.003002
282	-0.009162	-0.011203	-0.008327
283	-0.022810	-0.022710	-0.015569
284	0.011227	0.030783	0.025791
285	0.020256	0.007971	0.029153
286	0.006312	0.004402	-0.002592
287	0.004635	0.000270	0.004286
288	0.007169	0.005953	-0.003225
289	0.003925	-0.002481	0.004007
290	0.021658	0.018270	0.010885
291	0.015223	0.018352	0.011655
292	-0.039356	-0.025910	-0.023944
293	-0.014545	-0.005284	0.003789
294	-0.005625	0.000407	0.002024
295	-0.011869	-0.002596	-0.037638
296	0.004081	0.034111	0.032982
297	0.019457	0.023511	0.030694
298	-0.008466	0.005398	0.013120
299	0.012948	0.002470	0.010927
300	0.004622	-0.008716	-0.001185
301	0.003168	0.011972	0.004041
302	0.014112	0.018966	0.005419
303	-0.032688	-0.022239	-0.034072
304	-0.015854	-0.025912	-0.023414
305	-0.014307	-0.023744	-0.015440
306	-0.031565	-0.032543	-0.018246
307	-0.022926	-0.022732	-0.017643
308	-0.035830	-0.023075	-0.042604
309	-0.031859	-0.018211	-0.027860
310	0.011363	0.018872	0.032158
311	-0.017037	-0.009842	-0.008423
312	-0.003660	0.002425	0.004697
313	-0.004986	0.005376	0.005902
314	0.015454	0.014024	-0.005495
315	0.009450	0.017041	0.022944
316	0.004834	0.018513	0.016589
317	0.019387	0.031584	0.023265
318	-0.008927	-0.003330	-0.008234
319	0.017049	0.015804	0.028303
320	0.013157	0.011321	0.010970
321	0.002372	0.006000	0.011780
322	0.010288	0.016533	0.016426
323	-0.010442	-0.011450	-0.029115
324	-0.012699	-0.021046	-0.004936
325	-0.012361	-0.018238	-0.030915
326	-0.005537	-0.013913	-0.017504
327	-0.000691	-0.013755	-0.007317
328	0.004839	0.017618	0.014632
329	-0.005212	-0.012967	-0.009601
330	-0.010996	-0.003419	-0.007499
331	-0.016720	-0.020456	-0.016675
332	0.011342	-0.000870	-0.007356
333	-0.006232	0.004120	0.020070
334	-0.000680	-0.019593	0.000894
335	-0.020660	-0.027592	-0.033533
336	0.021052	0.002160	-0.008878
337	-0.008077	-0.003858	-0.009390
338	0.004805	0.007782	-0.011594
339	0.015360	0.020740	0.027269
340	0.011360	-0.007403	-0.001362
341	-0.009057	-0.011952	-0.027271
342	-0.016343	0.010115	0.017278
343	-0.007849	-0.012648	-0.003369
344	-0.006853	-0.013000	-0.008904
345	-0.000581	0.005977	-0.020303
346	-0.016374	-0.011372	-0.000266
347	0.027928	0.002030	-0.006165
348	0.018197	0.025485	0.036383
349	0.002911	0.007016	0.016213
350	-0.003023	-0.008022	-0.020660
351	-0.014198	0.007659	-0.006585
352	0.009663	0.006735	0.008102
353	0.026649	0.010516	0.003276
354	0.000970	0.005055	-0.002384
355	0.001385	0.002672	0.003068
356	-0.011271	0.006505	0.000713
357	-0.029285	-0.034045-	-0.045475
358	0.015315	0.017873	0.020914
359	-0.008642	-0.004291	-0.012885
360	-0.025975	-0.006681	-0.004023
361	0.002840	0.015817	0.035015
362	-0.000532	0.000335	0.016526
363	-0.009838	-0.020165	-0.035695

Appendix C: GAMMAs
NATIONAL ADULT LITERACY SURVEY
Older Adult Household Sample (Age 65 and Over)

	Prose	Document	Quantitative		Prose	Document	Quantitative
1	-0.696468	-0.324630	-0.824542	57	-0.024500	0.029898	-0.016409
2	0.123029	0.119880	0.147492	58	0.079442	0.029608	0.071669
3	-0.097580	-0.094474	-0.130137	59	-0.023310	-0.034135	-0.029227
4	-0.097283	-0.094735	-0.106092	60	0.044806	0.015788	0.037719
5	-0.052646	-0.050894	-0.050007	61	-0.019246	0.020493	-0.009001
6	-0.086046	-0.038686	-0.064589	62	-0.047609	-0.033720	-0.043382
7	-0.058029	0.035078	0.041684	63	-0.009610	-0.005079	0.021909
8	-0.030114	-0.042473	-0.038501	64	0.045159	0.033510	0.066901
9	0.000868	0.006002	0.010457	65	0.051897	0.036301	-0.000071
10	0.053496	0.065762	0.043253	66	0.026239	0.006546	0.055275
11	-0.071577	-0.036766	-0.047078	67	-0.000712	0.008784	-0.033081
12	0.014299	0.010358	-0.026751	68	-0.022585	-0.008307	0.045138
13	0.010380	0.033935	0.093111	69	0.023823	-0.007268	0.008051
14	-0.037928	-0.035350	-0.087950	70	-0.004194	0.009707	0.000582
15	0.002085	0.007125	-0.016354	71	-0.011271	-0.026241	-0.002707
16	0.034798	0.022530	0.060733	72	-0.011443	-0.001259	-0.024047
17	-0.035506	-0.000170	-0.025345	73	0.043073	0.004858	-0.011253
18	0.017213	0.034045	0.053061	74	0.020409	0.015351	0.001920
19	0.022131	-0.006640	0.008396	75	0.011240	-0.009299	-0.005710
20	0.048416	0.020109	0.011272	76	-0.010902	-0.008767	-0.017549
21	0.007442	0.043462	0.055416	77	0.047943	0.048694	0.044118
22	0.067563	0.024180	0.044035	78	-0.005911	-0.007945	-0.008279
23	0.049652	0.020645	0.022927	79	0.040516	0.072596	0.032181
24	-0.001022	-0.009701	0.000518	80	0.036900	0.020577	0.007312
25	-0.007572	-0.015753	-0.018993	81	-0.004027	-0.009460	-0.007478
26	0.037918	0.050174	0.004959	82	-0.005248	0.006252	0.013758
27	0.035524	0.002470	-0.020008	83	-0.067181	-0.026969	-0.056243
28	-0.088530	-0.058078	-0.043295	84	-0.030582	-0.001451	-0.006401
29	-0.041914	0.010054	-0.024781	85	-0.036545	0.010698	-0.033329
30	-0.005407	-0.029887	0.005143	86	0.032094	-0.012134	0.037562
31	-0.029926	-0.011225	0.008100	87	-0.005105	0.023298	0.041108
32	0.023444	0.026504	-0.006011	88	-0.035690	0.009070	-0.060221
33	-0.016532	0.002392	0.009618	89	-0.026805	-0.013154	0.003084
34	-0.017628	-0.022893	-0.002464	90	0.014591	-0.001149	-0.003837
35	-0.007781	0.010038	0.015425	91	-0.007511	0.000992	-0.034354
36	0.015747	-0.051567	-0.028405	92	0.023488	0.038423	0.057389
37	0.036843	0.007948	-0.045100	93	-0.079262	-0.030409	-0.043431
38	-0.014928	-0.025005	-0.030620	94	0.047112	-0.003613	0.029281
39	-0.026670	0.002288	0.026609	95	0.005730	-0.030713	-0.081626
40	-0.016271	0.015658	-0.010740	96	-0.045284	-0.045453	-0.023600
41	-0.033092	0.003272	-0.010303	97	-0.065135	0.000155	-0.032325
42	-0.054687	-0.010039	-0.081402	98	0.037463	0.019470	0.042835
43	0.019839	-0.000102	0.008872	99	-0.047693	-0.051052	-0.013302
44	-0.004266	0.033758	0.021862	100	-0.002558	0.022163	0.002897
45	-0.001320	0.001041	-0.007604	101	-0.000929	0.001428	0.015207
46	-0.028499	-0.027781	-0.036694	102	0.042060	0.034595	0.088993
47	-0.009114	-0.003394	-0.007231	103	-0.052925	-0.037937	-0.029515
48	-0.000016	-0.003060	-0.009959	104	0.005270	0.026438	-0.007026
49	-0.034910	0.001639	-0.043949	105	-0.033133	0.003822	-0.012218
50	-0.015900	0.019943	0.004212	106	0.017097	0.005463	0.046978
51	0.020977	-0.000606	0.001908	107	-0.004712	0.020729	0.012114
52	-0.043321	-0.006252	0.001379	108	0.040607	0.052989	0.052313
53	0.066659	-0.023256	-0.013089	109	-0.011361	0.042556	0.051671
54	-0.016781	0.026629	-0.009363	110	0.033800	0.072213	0.015122
55	0.018554	-0.003041	0.013472	111	0.049855	0.033013	0.066751
56	0.014425	-0.038030	-0.020161	112	0.028950	0.017404	-0.019285

	Prose	Document	Quantitative		Prose	Document	Quantitative
113	-0.028384	-0.041704	-0.042578	176	0.053517	0.068158	0.023497
114	-0.001625	0.034958	0.003599	177	0.027245	-0.030043	-0.031639
115	0.051202	0.065211	0.051029	178	-0.013117	-0.014364	0.015226
116	-0.012271	-0.028767	-0.020717	179	0.047136	0.064496	0.029324
117	-0.053872	-0.034091	0.003628	180	0.095186	0.081221	0.118004
118	0.056890	0.072086	0.080958	181	-0.089629	-0.039962	-0.036358
119	0.002490	0.024539	-0.005496	182	-0.081434	-0.127756	-0.020975
120	0.032645	0.014529	0.017348	183	-0.059728	0.038371	0.048911
121	-0.039396	-0.013852	0.016028	184	0.053915	-0.008981	0.022197
122	-0.102990	-0.007515	-0.022090	185	0.024570	0.026764	-0.032070
123	0.101599	0.002721	-0.012629	186	0.046124	0.025208	0.034829
124	0.005297	-0.022331	-0.087485	187	-0.025142	-0.030545	-0.000408
125	0.002534	0.032235	0.002393	188	0.058637	0.002408	-0.043807
126	0.024925	0.036739	0.004810	189	-0.095430	-0.127006	-0.008203
127	-0.018023	0.001981	-0.033553	190	-0.099259	-0.073843	-0.010435
128	-0.016540	0.010016	-0.040248	191	0.010319	-0.034114	0.025940
129	0.046602	0.044743	0.031324	192	0.078573	0.106708	0.087075
130	0.037049	0.034375	0.044014	193	-0.096024	-0.039775	-0.024144
131	-0.038138	-0.029644	-0.050953	194	0.056318	0.013941	0.063666
132	0.081140	0.080645	0.040575	195	0.041200	-0.001490	0.073673
133	0.041220	0.019105	0.035682	196	0.078400	0.064251	0.033923
134	0.062590	0.020505	-0.012509	197	0.010087	0.052857	0.032384
135	0.006133	0.036070	-0.004938	198	0.010744	-0.013520	0.031792
136	0.004026	0.015198	0.008078	199	-0.042792	-0.055403	0.001000
137	0.061381	-0.002229	0.038875	200	-0.010426	-0.031121	-0.012746
138	0.007543	-0.031943	-0.043280	201	0.037168	0.025947	-0.050658
139	-0.090683	-0.037671	-0.062426	202	0.076217	0.028951	0.049070
140	-0.005264	0.005673	0.024687	203	0.024920	-0.041832	-0.001721
141	-0.042486	-0.010730	0.007184	204	-0.114888	-0.047669	-0.094505
142	-0.023105	-0.003048	-0.035171	205	0.023248	0.022337	0.014508
143	-0.060511	-0.001009	0.054086	206	0.070822	0.001384	0.035751
144	0.008030	0.009280	-0.012930	207	0.005382	-0.009516	0.067701
145	-0.049440	-0.052883	-0.038357	208	-0.010340	-0.058793	0.009885
146	-0.000030	-0.019706	-0.002429	209	-0.017292	0.021813	-0.021749
147	0.016900	-0.006279	0.044934	210	0.007379	-0.020276	-0.050906
148	0.020380	0.009979	0.007967	211	0.076506	0.080075	0.057274
149	-0.038262	0.007680	-0.047827	212	-0.068471	0.086012	-0.020107
150	0.071183	0.086266	0.027368	213	-0.010257	-0.044884	-0.009749
151	-0.041660	0.012230	-0.021877	214	-0.085220	-0.039352	-0.114865
152	0.032483	-0.027630	-0.013635	215	0.020931	0.016353	0.090330
153	-0.052609	0.002632	-0.009997	216	-0.007994	-0.008726	0.089925
154	0.037496	-0.003962	0.012940	217	-0.013781	0.014194	0.011713
155	-0.041408	-0.061123	-0.031518	218	0.046753	0.003008	0.033913
156	-0.003278	-0.022905	-0.019210	219	0.036094	0.024667	-0.010037
157	-0.029915	-0.040653	0.000886	220	-0.067357	-0.055975	-0.110139
158	-0.149524	-0.046008	-0.052838	221	-0.082353	-0.024666	-0.012930
159	-0.034975	-0.018606	0.033034	222	0.007705	0.000582	0.007445
160	-0.082961	-0.068835	-0.076822	223	0.014729	0.035137	0.033280
161	-0.123266	-0.051080	0.019153	224	-0.032909	-0.054567	-0.040668
162	-0.040868	-0.014158	-0.008891	225	-0.057445	-0.035665	-0.051152
163	-0.000118	0.055851	0.043320	226	-0.012623	-0.006733	-0.061505
164	-0.131761	-0.091515	-0.079388	227	0.010394	-0.023002	0.057118
165	-0.125266	-0.097010	-0.020293	228	-0.036654	0.020282	-0.023880
166	0.120325	0.122827	0.021676	229	0.052759	-0.001380	0.042809
167	-0.095472	-0.068295	0.000718	230	-0.062926	-0.053527	-0.062693
168	-0.060790	0.001388	0.010366	231	0.085731	0.065321	0.053125
169	-0.053578	-0.071990	0.014592	232	0.016716	0.016639	0.007850
170	0.093001	0.034875	-0.016046	233	0.023704	-0.034080	-0.057031
171	0.067564	0.050581	0.025562	234	-0.071324	-0.023398	-0.060009
172	0.008479	-0.044934	-0.033439	235	0.054490	0.015105	-0.043954
173	0.037565	0.037211	0.115551	236	0.025261	0.059712	0.092530
174	0.087492	0.103395	0.068621	237	-0.043365	0.008669	-0.016895
175	-0.100635	-0.011538	-0.066955	238	-0.046877	-0.004387	-0.034713

	Prose	Document	Quantitative
239	0.074573	-0.025891	0.009077
240	-0.070373	-0.017536	-0.090507
241	0.041324	0.075884	-0.013010
242	0.020968	0.056989	0.082148
243	0.003493	0.036686	0.006098
244	-0.136340	-0.087663	-0.068879
245	0.034485	-0.076457	-0.106375
246	-0.059404	0.041748	0.090852
247	0.014792	-0.040449	0.024501
248	0.024626	0.035784	0.027012
249	0.002038	0.009055	-0.021120
250	0.011074	-0.006154	-0.008240
251	0.010565	0.002191	0.005604
252	-0.001337	-0.000135	0.000067
253	-0.015997	-0.018208	-0.065155
254	0.031884	0.019099	-0.006782
255	0.022806	0.010832	0.018750
256	-0.047281	-0.012296	0.018907
257	-0.058781	0.006861	-0.027564
258	-0.001453	-0.025525	-0.128271
259	-0.015247	-0.052329	-0.020261
260	-0.114785	-0.095892	-0.006029
261	-0.031691	0.011521	0.045710
262	0.075018	0.014134	-0.077921
263	0.013794	-0.004485	0.007900
264	-0.024124	-0.057904	0.007305
265	-0.003682	-0.017131	0.030737
266	0.020039	-0.011149	0.016667
267	0.029544	0.032481	0.003408
268	-0.051239	-0.038409	0.016057
269	0.007642	-0.021031	-0.082381
270	0.019068	-0.041019	-0.055735
271	0.060972	0.063323	0.009870
272	0.020789	-0.016149	-0.020259
273	0.034973	0.022547	-0.057658
274	-0.032169	-0.060975	-0.054036
275	0.073935	0.010137	0.064657
276	-0.023246	0.012239	-0.010071
277	0.005462	0.003324	-0.021008
278	-0.093062	-0.003271	-0.013992
279	0.013211	0.016054	-0.000961
280	0.011788	-0.016551	-0.005429
281	0.017884	0.012250	0.036243
282	-0.006219	-0.018212	-0.009014
283	-0.042659	0.015312	-0.003625
284	-0.032679	0.025279	0.028521
285	0.034593	0.060918	0.060077
286	0.045853	-0.008138	0.098533
287	-0.083068	-0.044243	-0.074861
288	0.017543	-0.006686	0.079878
289	-0.032950	-0.012982	-0.014884
290	0.003796	0.036176	0.013545
291	-0.048610	-0.065322	-0.061833
292	-0.052734	-0.021560	-0.006262
293	-0.019607	-0.080347	-0.029365
294	-0.029012	-0.021472	-0.014167
295	-0.004429	0.094150	0.107406
296	0.013387	0.001851	-0.040036
297	-0.050987	-0.082685	-0.055733
298	-0.004564	0.021338	0.069134
299	-0.062488	0.013864	-0.065395
300	-0.053988	-0.058201	-0.033302
301	-0.034463	0.023448	-0.018039
302	-0.059383	-0.001880	0.000043

Appendix C: GAMMAs
NATIONAL ADULT LITERACY SURVEY
Prison Sample

	Prose	Document	Quantitative		Prose	Document	Quantitative
1	-0.386127	0.115492	-0.679943	56	0.025622	0.006128	0.009158
2	0.098797	0.115528	0.084354	57	0.008399	0.013486	0.003544
3	-0.119333	-0.147741	-0.112306	58	0.023643	0.013188	0.053960
4	-0.009922	-0.008031	-0.033271	59	-0.007247	-0.015498	-0.008535
5	0.038080	0.069555	0.064672	60	0.008813	0.036166	0.046600
6	-0.101066	-0.101378	-0.128542	61	0.011741	-0.028406	-0.047608
7	-0.006638	0.002384	-0.006024	62	0.073372	0.015817	0.071543
8	0.045433	0.088297	0.056090	63	0.012914	0.014476	0.018365
9	0.063591	0.021258	0.042825	64	-0.008764	-0.010450	-0.023790
10	-0.010620	-0.020706	-0.019262	65	0.026168	0.010661	0.029063
11	-0.071385	0.000707	-0.007395	66	-0.036864	-0.010824	-0.059075
12	0.072714	0.060405	0.056821	67	-0.023476	-0.019117	-0.011770
13	0.002147	0.003294	0.008546	68	0.012919	-0.012260	-0.026808
14	0.039614	-0.016900	0.013515	69	0.034293	0.063347	0.046600
15	-0.043960	-0.047101	-0.048928	70	0.059590	0.067941	0.041859
16	0.015271	0.005392	-0.010340	71	0.010522	0.022620	0.036283
17	0.050484	0.052416	0.051440	72	0.046969	0.028438	0.039097
18	0.009645	-0.000991	0.037162	73	-0.016075	0.008363	-0.031558
19	-0.029589	0.004990	-0.005317	74	0.022733	-0.005260	0.045034
20	0.020002	0.027607	0.063484	75	0.048615	0.019959	0.013867
21	0.003605	-0.001052	0.013750	76	0.022498	0.004851	0.036237
22	0.050437	0.025307	0.059811	77	0.041916	0.048629	0.020832
23	0.025273	0.066893	0.078071	78	0.019887	-0.005369	-0.008488
24	0.015111	0.026809	0.036992	79	-0.038142	-0.045681	-0.013027
25	0.008746	0.021444	0.058359	80	-0.008103	-0.049953	-0.005571
26	0.009600	0.015503	0.048090	81	-0.021502	0.042218	0.024686
27	0.016856	0.017308	-0.000435	82	-0.009873	0.014575	-0.054513
28	-0.062733	-0.054986	-0.033628	83	-0.003567	0.009675	-0.057259
29	0.029458	-0.013786	0.035065	84	-0.015850	-0.011504	-0.019602
30	-0.011410	-0.038673	-0.015364	85	0.010857	-0.016347	-0.043510
31	0.018421	0.004432	0.010688	86	-0.012352	-0.008561	-0.037641
32	0.004188	0.001543	0.010891	87	0.031373	0.020015	0.011665
33	-0.022714	-0.029624	0.012059	88	0.011445	0.003191	0.008798
34	0.043681	0.034507	0.011151	89	-0.039701	-0.011745	-0.006750
35	0.029949	0.000454	0.023246	90	-0.003997	-0.014574	-0.001786
36	-0.013308	0.028778	-0.018692	91	-0.055559	-0.011462	-0.064019
37	0.001964	-0.024369	0.051740	92	0.060769	0.015116	0.038254
38	-0.001353	-0.016757	-0.019114	93	0.014078	0.040705	0.025106
39	-0.021444	-0.053042	-0.026912	94	0.017253	0.032904	0.027442
40	0.009998	0.014635	0.021932	95	0.017401	0.024886	0.013946
41	0.015156	0.004046	0.021452	96	0.023030	0.045523	0.065414
42	0.018623	0.049478	0.030782	97	-0.004253	0.022086	0.014789
43	0.023381	-0.012094	0.018388	98	0.055388	0.068430	0.105604
44	0.015585	0.006362	-0.001647	99	0.053841	0.019025	-0.008768
45	0.014962	0.003746	-0.023795	100	-0.053040	-0.019697	-0.055662
46	0.010497	0.006209	0.020963	101	-0.009601	-0.003790	0.016514
47	-0.047161	-0.045674	-0.051391	102	0.008010	0.023432	0.042046
48	-0.008596	-0.025872	0.048486	103	0.004746	-0.021115	0.005169
49	0.007507	0.048285	0.041077	104	0.002392	0.014636	0.047088
50	-0.001013	0.019335	-0.015324	105	0.002100	-0.034194	-0.039841
51	0.047688	0.070190	0.097123	106	-0.016487	-0.027438	-0.043442
52	-0.010818	-0.006446	0.043799	107	-0.023802	0.010130	-0.029880
53	0.031026	0.034537	0.047204	108	0.004812	-0.019412	0.051052
54	-0.010685	-0.045518	-0.040707	109	-0.032711	-0.046812	-0.051037
55	0.015535	-0.010227	-0.004162	110	0.013359	0.049220	0.061664

	Prose	Document	Quantitative		Prose	Document	Quantitative
111	-0.037045	-0.006881	0.029014	174	-0.044997	-0.007263	-0.031639
112	0.018897	0.010860	0.036167	175	0.006280	-0.003237	-0.038459
113	-0.025220	-0.010970	-0.027394	176	-0.016205	0.026528	0.057793
114	-0.028522	-0.032548	0.007457	177	-0.045052	-0.027313	-0.072989
115	-0.048154	-0.068954	-0.074520	178	-0.023367	0.045511	0.042296
116	-0.035048	-0.029749	-0.022708	179	0.020901	0.058120	0.035285
117	-0.018531	-0.049987	-0.059223	180	-0.051920	-0.039995	-0.032435
118	0.003765	0.011666	-0.002760	181	0.044260	0.037946	0.041944
119	0.043675	-0.030724	-0.011411	182	-0.014558	-0.015115	0.012966
120	0.017049	0.031123	0.036281	183	0.005376	-0.000464	0.021033
121	-0.042115	-0.018285	-0.032984	184	-0.010394	-0.013561	-0.098017
122	0.002604	-0.026589	-0.028350	185	0.000516	0.033379	-0.016930
123	0.020407	0.049889	0.036644	186	0.033231	0.020421	0.053701
124	-0.044875	-0.042949	-0.004192	187	0.027038	0.034328	-0.035205
125	0.032112	0.021865	0.051860	188	0.020023	0.039011	0.027479
126	0.017408	-0.017913	0.029637	189	0.045812	0.010914	0.053145
127	0.015177	0.064081	-0.012195	190	-0.055033	-0.022511	-0.028896
128	-0.014977	-0.006745	0.000607	191	0.044691	-0.023748	0.050027
129	-0.015005	-0.024953	-0.036312	192	0.010190	0.006443	-0.017211
130	0.004399	0.029152	0.009712	193	-0.021526	-0.023708	-0.029452
131	0.008009	0.011945	0.018685	194	0.030060	0.007809	0.025324
132	-0.037840	-0.022206	-0.017948	195	0.032094	-0.004360	0.024545
133	-0.029476	-0.021716	0.025163	196	-0.006352	0.057813	-0.026279
134	0.032954	-0.004876	0.041474	197	0.049916	0.019031	0.073415
135	0.001990	0.001866	-0.027223	198	0.029459	0.003176	-0.030769
136	-0.007407	0.001661	-0.000307	199	-0.058354	-0.046693	-0.098602
137	-0.003584	0.043374	0.026743	200	-0.056031	-0.025985	-0.003133
138	0.011257	0.016954	-0.048631	201	0.004596	-0.032438	-0.092215
139	-0.027242	-0.040754	-0.053129	202	0.045533	-0.022636	0.000578
140	0.038098	-0.014581	-0.011786	203	0.009320	-0.010225	0.006611
141	-0.021438	-0.000328	-0.041456	204	-0.052631	-0.050596	0.012799
142	0.013231	-0.001864	-0.015223	205	-0.031363	-0.001457	-0.068431
143	0.021533	-0.005629	-0.015153	206	0.034183	0.046132	0.028984
144	-0.022164	0.027300	0.016554	207	-0.029619	0.034883	-0.024560
145	0.028294	0.036726	0.046019	208	0.023537	0.023194	0.007812
146	0.015652	0.035682	-0.017865	209	-0.034443	-0.004238	-0.039429
147	0.018223	-0.001434	-0.019782	210	-0.001119	0.028951	-0.026840
148	-0.001679	-0.027476	-0.017189	211	-0.040295	-0.003369	-0.036675
149	0.030393	0.022408	-0.000927	212	-0.012652	-0.018627	0.009679
150	0.013463	-0.028601	-0.091329	213	-0.014955	-0.004247	0.019465
151	0.026264	0.056480	0.033862	214	0.015441	0.004260	0.005536
152	0.002077	-0.021658	-0.040907	215	0.016017	0.006281	-0.023654
153	0.000157	0.038946	0.030253	216	-0.000645	0.044151	0.064389
154	0.002772	-0.003294	0.002876	217	-0.064714	0.040988	0.020513
155	-0.021192	-0.038986	-0.059001	218	0.021936	-0.025010	0.010566
156	0.022895	-0.048764	0.001878	219	0.025180	0.053746	0.041178
157	0.015804	0.032472	0.067740	220	0.052034	0.061584	0.092301
158	-0.060983	-0.031499	-0.064658	221	-0.001335	0.035925	0.039059
159	-0.013410	-0.008367	0.015238	222	0.007884	0.013447	0.041253
160	-0.007173	-0.032191	-0.074340	223	0.036239	-0.025363	-0.007983
161	0.044998	0.014985	0.001677	224	-0.056438	-0.059634	0.009229
162	0.018962	0.005356	0.014188	225	-0.040164	-0.035788	-0.029660
163	-0.022333	-0.003362	-0.015549	226	-0.032352	-0.019189	-0.050331
164	-0.041969	-0.038665	-0.031970	227	0.012939	0.032009	0.015540
165	-0.022325	0.015666	-0.001232	228	0.018708	0.013918	0.006742
166	0.017775	0.026726	0.012530	229	-0.068357	-0.080536	-0.103954
167	-0.004095	0.022341	-0.016214	230	-0.048546	-0.012633	-0.004435
168	0.028579	-0.020722	0.001295	231	0.033861	0.044524	0.058598
169	-0.002153	-0.013035	-0.029983	232	0.044909	0.026979	-0.003536
170	0.066341	0.032730	0.035171	233	0.080741	0.122653	0.082427
171	-0.005796	-0.004321	-0.012371	234	-0.004219	0.013281	-0.031707
172	0.001836	0.005711	0.029768	235	0.011301	0.050581	0.052976
173	0.005651	0.000893	0.021285	236	-0.035830	-0.005951	0.024603

	Prose	Document	Quantitative		Prose	Document	Quantitative
237	0.039636	-0.008417	-0.012227				
238	0.047586	0.066837	0.049539				
239	0.039380	0.022710	-0.067035				
240	0.003391	-0.018961	-0.019717				
241	-0.012766	-0.037832	-0.032727				
242	-0.014758	-0.033731	-0.035291				

Appendix D

RP80s AND ITEM PROBABILITIES X 100

Appendix D: RP80s and Item Probabilities X 100
NATIONAL ADULT LITERACY SURVEY

PROSE								
Item #	RP80	150	200	250	300	350	400	450
AB21101	207.9	31.	75.	95.	99.	100.	100.	100.
AB21201	250.0	10.	40.	80.	96.	99.	100.	100.
AB30501	370.3	5.	12.	27.	50.	73.	88.	95.
AB30601	332.6	2.	7.	24.	60.	87.	97.	99.
AB31201	362.3	2.	7.	21.	47.	75.	91.	97.
AB40901	328.1	3.	11.	32.	65.	88.	97.	99.
AB41001	262.0	28.	53.	76.	90.	96.	99.	100.
AB50101	467.6	3.	6.	12.	23.	39.	58.	75.
AB50201	342.2	20.	21.	26.	50.	84.	97.	100.
AB60201	280.0	2.	13.	54.	90.	99.	100.	100.
AB60601	304.1	6.	20.	48.	78.	93.	98.	100.
AB70401	288.6	13.	33.	62.	84.	94.	98.	99.
AB71001	274.8	27.	33.	61.	92.	99.	100.	100.
AB71101	347.3	2.	8.	24.	54.	81.	94.	98.
NC00301	149.0	81.	95.	99.	100.	100.	100.	100.
NC00401	223.6	38.	69.	89.	97.	99.	100.	100.
N010101	348.5	1.	5.	19.	50.	81.	95.	99.
N010201	287.6	3.	15.	52.	86.	97.	100.	100.
N010301	432.5	0.	1.	3.	11.	32.	63.	86.
N080101	225.7	12.	56.	92.	99.	100.	100.	100.
N080201	276.5	1.	8.	51.	93.	99.	100.	100.
N080301	361.6	5.	13.	29.	53.	76.	90.	96.
N090601	252.7	1.	13.	77.	99.	100.	100.	100.
N090701	256.1	1.	12.	74.	98.	100.	100.	100.
N090801	359.1	0.	1.	4.	26.	73.	96.	99.
N100101	238.3	5.	38.	88.	99.	100.	100.	100.
N100201	283.4	1.	10.	49.	89.	99.	100.	100.
N100301	286.5	2.	12.	49.	87.	98.	100.	100.
N100401	382.5	1.	2.	9.	28.	62.	87.	96.
N110101	261.1	9.	35.	74.	94.	99.	100.	100.
N110401	313.3	6.	18.	44.	74.	91.	97.	99.
N110501	275.7	7.	27.	64.	90.	98.	100.	100.

PROSE								
Item #	**RP80**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
N110601	410.0	0.	0.	2.	8.	33.	74.	94.
N120301	315.2	1.	6.	28.	70.	93.	99.	100.
N120401	316.8	1.	4.	23.	68.	94.	99.	100.
N120501	423.0	0.	0.	2.	8.	30.	66.	90.
N120901	209.5	41.	75.	93.	98.	100.	100.	100.
N130201	323.7	1.	4.	22.	63.	91.	99.	100.
N130301	374.4	0.	1.	7.	26.	64.	90.	98.
N130401	346.1	0.	0.	3.	26.	83.	99.	100.
N130801	273.5	16.	40.	69.	89.	96.	99.	100.

DOCUMENT								
Item #	**RP80**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
SCOR100	69.2	94.	97.	99.	100.	100.	100.	100.
SCOR300	151.1	80.	95.	99.	100.	100.	100.	100.
AB20101	255.1	27.	43.	77.	95.	99.	100.	100.
AB20201	277.3	16.	32.	64.	89.	97.	99.	100.
AB20301	306.7	15.	20.	40.	76.	95.	99.	100.
AB20401	288.6	38.	53.	69.	83.	91.	96.	98.
AB20501	290.4	40.	55.	70.	82.	90.	95.	97.
AB20601	265.2	16.	34.	71.	93.	99.	100.	100.
AB20701	323.5	26.	38.	56.	73.	86.	94.	97.
AB20801	313.1	15.	17.	31.	70.	95.	99.	100.
AB20901	347.5	16.	17.	22.	45.	81.	97.	100.
AB21001	314.4	20.	29.	50.	74.	90.	97.	99.
AB21501	260.4	18.	45.	75.	92.	98.	99.	100.
AB30101	231.7	24.	60.	88.	97.	99.	100.	100.
AB30301	239.9	22.	55.	84.	96.	99.	100.	100.
AB30401	296.3	14.	33.	59.	81.	93.	98.	99.
AB30701	244.7	17.	50.	83.	96.	99.	100.	100.
AB30801	346.3	4.	12.	30.	58.	81.	93.	98.
AB30901	378.7	30.	41.	53.	65.	75.	83.	89.
AB31001	303.3	6.	19.	48.	79.	94.	98.	100.
AB31101	279.5	14.	35.	66.	87.	96.	99.	100.
AB31301	234.7	35.	64.	85.	95.	98.	100.	100.
AB40101	233.2	30.	62.	86.	96.	99.	100.	100.
AB40401	313.3	1.	6.	29.	71.	94.	99.	100.
AB50401	245.9	26.	56.	82.	94.	98.	100.	100.
AB50402	313.1	6.	19.	45.	74.	91.	97.	99.
AB50501	328.2	33.	47.	61.	74.	84.	90.	94.
AB50601	275.4	6.	25.	64.	90.	98.	100.	100.
AB50701	253.7	7.	33.	78.	96.	99.	100.	100.
AB50801	223.5	41.	70.	88.	96.	99.	100.	100.
AB50901	274.9	10.	31.	66.	89	97.	99.	100.
AB60101	245.1	7.	39.	83.	97.	100.	100.	100.

DOCUMENT								
Item #	**RP80**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
AB60102	245.2	5.	33.	83.	98.	100.	100.	100.
AB60103	244.7	16.	48.	83.	96.	99.	100.	100.
AB60104	237.5	5.	38.	88.	99.	100.	100.	100.
AB60301	199.3	28.	81.	98.	100.	100.	100.	100.
AB60302	201.8	39.	79.	96.	99.	100.	100.	100.
AB60303	248.6	18.	49.	81.	95.	99.	100.	100.
AB60304	309.0	4.	14.	42.	76.	93.	98.	100.
AB60305	291.9	17.	37.	62.	83.	93.	98.	99.
AB60306	236.6	21.	56.	86.	97.	99.	100.	100.
AB60501	377.5	0.	1.	4.	20.	60.	90.	98.
AB60502	317.7	1.	6.	27.	68.	93.	99.	100.
AB60701	205.4	32.	77.	96.	99.	100.	100.	100.
AB60801	198.3	44.	81.	96.	99.	100.	100.	100.
AB60802	186.8	48.	87.	98.	100.	100.	100.	100.
AB60803	180.3	49.	91.	99.	100.	100.	100.	100.
AB61001	216.4	43.	73.	90.	97.	99.	100.	100.
AB70104	192.8	65.	82.	92.	96.	99.	99.	100.
AB70301	286.6	12.	32.	62.	85.	95.	99.	100.
AB70701	283.8	11.	31.	62.	86.	96.	99.	100.
AB70801	228.3	18.	58.	90.	98.	100.	100.	100.
AB70901	223.1	27.	65.	91.	98.	100.	100.	100.
AB70902	270.8	12.	35.	69.	90.	97.	99.	100.
AB70903	222.6	16.	61.	93.	99.	100.	100.	100.
AB71201	188.0	55.	85.	96.	99.	100.	100.	100.
N010401	214.0	38.	73.	92.	98.	100.	100.	100.
N010801	265.5	16.	41.	73.	91.	97.	99.	100.
N010901	385.5	2.	5.	14.	35.	63.	85.	95.
N011001	316.6	11.	25.	49.	74.	89.	96.	99.
N080601	304.7	2.	10.	38.	77.	95.	99.	100.
N080701	259.7	7.	32.	74.	94.	99.	100.	100.
N080801	322.8	5.	16.	39.	69.	89.	97.	99.
N080802	232.8	9.	48.	90.	99.	100.	100.	100.
N090301	210.9	30.	73.	94.	99.	100.	100.	100.

DOCUMENT								
Item #	RP80	150	200	250	300	350	400	450
N090401	256.8	11.	39.	76.	94.	99.	100.	100.
N090501	284.7	17.	38.	65.	85.	94.	98.	99.
N100501	303.7	2.	11.	39.	78.	95.	99.	100.
N100601	342.2	0.	2.	12.	46.	84.	97.	100.
N100701	395.7	0.	0.	2.	10.	43.	82.	97.
N110301	250.6	22.	51.	80.	94.	98.	100.	100.
N110302	242.9	31.	59.	83.	94.	98.	99.	100.
N110701	334.2	19.	34.	52.	70.	84.	92.	96.
N110901	387.3	1.	2.	9.	27.	59.	85.	96.
N120101	238.4	19.	54.	85.	97.	99.	100.	100.
N120201	266.6	14.	39.	72.	91.	98.	99.	100.
N120601	212.6	44.	74.	91.	98.	99.	100.	100.
N130101	258.6	1.	15.	72.	97.	100.	100.	100.
N130102	242.2	8.	41.	85.	98.	100.	100.	100.
N130103	242.3	1.	18.	87.	100.	100.	100.	100.
N130104	251.2	0.	10.	79.	99.	100.	100.	100.

QUANTITATIVE								
Item #	**RP80**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
AB40201	355.3	2.	7.	21.	49.	78.	93.	98.
AB40301	310.6	46.	50.	60.	76.	91.	97.	99.
AB40501	325.4	3.	10.	31.	66.	89.	97.	99.
AB40601	294.5	5.	19.	51.	83.	96.	99.	100.
AB40701	277.1	1.	8.	51.	93.	99.	100.	100.
AB40702	280.8	0.	3.	38.	93.	100.	100.	100.
AB40703	282.2	0.	3.	37.	92.	100.	100.	100.
AB40704	283.1	0.	2.	34.	92.	100.	100.	100.
AB50301	432.6	2.	5.	11.	24.	45.	68.	85.
AB50403	381.5	4.	11.	24.	45.	69.	85.	94.
AB50404	270.4	9.	32.	68.	91.	98.	100.	100.
AB60901	344.5	15.	29.	47.	66.	81.	91.	96.
AB70501	330.9	2.	9.	29.	63.	87.	97.	99.
AB70601	348.0	2.	6.	21.	52.	81.	94.	99.
AB70904	220.1	37.	70.	90.	97.	99.	100.	100.
NC00501	191.3	63.	83.	93.	98.	99.	100.	100.
NC00601	246.0	31.	59.	81.	93.	98.	99.	100.
N010501	286.6	9.	28.	60.	85.	96.	99.	100.
N010601	355.0	0.	2.	9.	36.	77.	95.	99.
N010701	341.7	1.	4.	17.	51.	84.	96.	99.
N011101	329.9	1.	5.	23.	60.	89.	98.	100.
N080401	368.2	3.	9.	23.	47.	73.	89.	96.
N080501	321.3	6.	18.	42.	71.	89.	96.	99.
N080901	375.0	1.	4.	12.	35.	67.	89.	97.
N081001	343.4	4.	13.	32.	59.	82.	94.	98.
N090101	308.4	1.	4.	25.	74.	96.	100.	100.
N090201	284.0	0.	4.	40.	90.	99.	100.	100.
N090901	277.8	14.	37.	67.	87.	96.	99.	100.
N091001	420.6	2.	5.	12.	26.	49.	73.	88.
N100801	250.0	34.	59.	80.	92.	97.	99.	100.
N100901	333.0	10.	23.	44.	68.	85.	94.	98.
N101001	277.8	8.	28.	64.	89.	97.	99.	100.

QUANTITATIVE								
Item #	**RP80**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
N110201	359.6	0.	2.	10.	36.	75.	94.	99.
N110303	238.3	31.	61.	84.	95.	99.	100.	100.
N110801	299.7	6.	20.	50.	80.	94.	99.	100.
N120701	279.1	9.	29.	63.	88.	97.	99.	100.
N120801	365.3	0.	2.	9.	33.	71.	92.	98.
N121001	331.8	1.	5.	22.	59.	88.	97.	99.
N121101	405.1	0.	1.	4.	14.	43.	77.	94.
N130501	388.4	3.	8.	19.	39.	64.	84.	94.
N130601	306.3	2.	9.	36.	76.	95.	99.	100.
N130701	435.5	0.	1.	3.	10.	29.	61.	86.
N130901	350.4	1.	4.	17.	47.	80.	95.	99.

Appendix E

NON-INTERVIEW REPORT FORM

Appendix E: Non-Interview Report Form
NATIONAL ADULT LITERACY SURVEY

This section contains excerpts from the interviewers' manual describing how the Non-interview Report form was to be completed and the interviewers' observations on how the respondent completed the exercises. Responses to these questions are included in the data file.

Whenever you are unable to complete a Background Questionnaire and/or an Exercise Booklet with the respondent, you must complete the Non-interview Report (NIR) form. This form is a separate, single page (double-sided) document that can easily be identified by its purple markings along the sides of the page. It should also be filled out when you have partially completed either the Background Questionnaire or Exercise Booklet. That is, you would complete the NIR when the respondent broke off the interview or exercise before you were able to ask all applicable questions in the Background questionnaire or lead the respondent through all of the tasks in the Exercise Booklet.

Be sure to write your name and the date of the non-response in the spaces provided at the top of side one on the NIR form. At the top of the page on side one, you will also notice space for a Westat ID label and an ID number grid box. You will be responsible for attaching a mini-label, writing in the ID number (including the respondent number) at the top of the grid box from the label, as well as filling in the corresponding bubbles in the box for the ID. In addition to the Westat ID number, if you began the Exercise with the respondent, but are filling out the NIR because the respondent was not able to complete the Exercise Booklet, please write in the Booklet ID number in the spaces provided at the top of the page and fill in the corresponding bubbles. Please be aware that side one of the NIR form is a "scannable" document and side two is not. Please write any additional comments you may have concerning the nonresponse on page two only, leaving side one free of stray marks and marginal notes.

BACKGROUND QUESTIONNAIRE
AND EXERCISE NON-INTERVIEW REPORT (NIR)

NAME OF INTERVIEWER: ______________________________ DATE: ________________

CAREFULLY PLACE LABEL INSIDE THIS BOX

⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	
①	①	①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②	②	②
③	③	③	③	③	③	③	③	③	③	
④	④	④	④	④	④	④	④	④	④	
⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	
⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	
⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	
⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	
⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	

ENTER BOOKLET NUMBER

⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①
②	②	②	②	②	②
③	③	③	③	③	③
④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨

1. What is the status of this case?

BQ nonresponse: respondent did not begin BQ .... ①
BQ nonresponse: respondent started but did not complete BQ .... ②
EX nonresponse: respondent completed BQ but did not start Exercise .... ③
EX nonresponse: respondent completed BQ and started but did not complete Exercise
(Enter booklet ID number above) .... ④

2. Why were you unable to begin or complete the Background Questionnaire or Exercise?

Non-English language .... ① **(Q3)**
Physical or mental disability .... ② **(Q4)**
Reading and/or writing difficulty .... ③ **(Q5)**
Respondent refused .... ④ **(Q6)**
Someone refused for respondent .... ⑤ **(Q6)**
Maximum calls .... ⑥ **(Q7)**
Unavailable during field period .... ⑦ **(Q7)**
Other reason .... ⑧ **(Q7)**

3. Which non-English language did the respondent speak, read, and/or write?

Spanish .... ① **(BOX 1)**
Other language (SPECIFY): ______________ .... ② } **(END OF NIR)**
Could not determine .... ③ } **(END OF NIR)**

BOX 1. If this is a BQ nonresponse (Q.1 coded 1 or 2), go to Q.8. Otherwise, end of NIR.

4. Code the nature of respondent's disability. Code all that apply.

Learning disability .... ①
Any mental or emotional condition .... ②
Mental retardation .... ③
Hearing impairment .... ④
Blind/Visual impairment .... ⑤
Speech/Language impairment .... ⑥
A physical disability .... ⑦
Other (SPECIFY): ______________ ⑧

BOX 2. If this is a BQ nonresponse (Q.1 coded 1 or 2), go to Q.8. Otherwise, end of NIR.

5. Explain in detail the nature of the respondent's reading and/or writing problem.

____

____

____

END OF NIR

6. Explain the circumstances surrounding the refusal or breakoff, recording what the respondent (or individual refusing for the respondent) said.

____

____

____

GO TO Q.8

7. Explain in detail why your were unable to begin or complete the Background Questionnaire or Exercise.

____

____

____

8. Record any information below that might help another interviewer complete the Background Questionnaire and/or the Exercise Booklet with this respondent.

____

____

____

9. Any additional comments.

____

____

____

Copyright © 1991 by Educational Testing Service Princeton, NJ 08541. All rights reserved.

Q.1 This item tells us whether this is a Background Questionnaire NIR or an Exercise NIR. It also notes the status of the BQ and/or Exercise. To code this question, follow these rules using the result code you have entered on the Record of Actions for the Background Questionnaire and Exercise:

- Code "BQ nonresponse: respondent did not begin BQ" if the result code for the Background Questionnaire is blank, 3 through 9, or 0.
- Code "BQ nonresponse: respondent started but did not complete BQ" if the result code for the Background Questionnaire is 2.
- Code "EX nonresponse: respondent completed BQ but did not start Exercise" if the result code for the Background Questionnaire is 1 and the result code for the Exercise is 07 through 11.
- Code "EX nonresponse: respondent completed BQ and started but did not complete Exercise" if the result code for the Background Questionnaire is 1 and the result code for the Exercise is 02 through 06.

Q.2 Your answer to this question should be consistent with the final result code you assigned at either the Background Questionnaire level or the Exercise level as the reason for the nonresponse. Refer to the following for definitions of the answer categories.

1 Non-English language - The respondent was unable to speak or read English well enough to begin or to complete the Background Questionnaire or Exercise. (If the respondent's non-English language is Spanish, you should have made an effort to use an adult to assist with the Spanish version of the Background Questionnaire.)

In determining when to assign this code, remember that there may be some respondents who cannot speak or understand English when spoken to but who are able to read in English. If a R is unable to speak or understand English, our rules state that he/she is unable to participate in the Background Questionnaire interview. However, if they can read English, they can complete the Exercise. If a non-English speaking respondent is able to complete the exercise, then try to use a translator to administer the BQ in the respondent's non-English language. This is the ONLY situation where a translator can be used to obtain a BQ interview in a language other than English or Spanish.

2 Physical or mental disability - The respondent was unable to complete or to begin the Background Questionnaire or Exercise due to a physical or mental impairment, such as a chronic illness, vision problems, hearing problems, or a psychological disorder.

3 Reading and/or writing difficulty - The respondent was unable to begin or to complete the Exercise due to insufficient reading and/or writing skills. This category should only include those respondents who cannot read or write for reasons other than a physical or mental disability.

4 Respondent Refused - The respondent was unwilling to begin or to complete the Background Questionnaire or Exercise.

5 Someone refused for respondent - An adult, not necessarily a household member, refused to allow the respondent to begin or to continue with the Background Questionnaire or Exercise. (Most commonly, this will be parents refusing for teenagers, or adults refusing for their elderly parents.)

6 Maximum calls - You were unable to complete the Background Questionnaire and Exercise after making at least four attempts (in addition to the Screener attempts) on different days and at different times.

7 Unavailable during field period - The selected respondent was out of town or was otherwise not available at anytime during the field period. This includes non-chronic illness that prevents participation for the duration of the field period.

8 Other reason - The respondent did not begin or complete the Background Questionnaire or Exercise for any other reason not covered elsewhere in this code structure.

Q.3 If you did not begin or complete the Background Questionnaire or Exercise because of a non-English language problem, that is, the respondent did not speak, read, or write English, code the respondent's non-English language at this question. (If the language is Spanish, you should have made an effort to use an adult to administer the Spanish version of the Background Questionnaire. Remember, however, a household translator cannot be used for any language other than Spanish.)

Q.4 If the respondent is unable to begin or to complete the Background Questionnaire or Exercise because of a physical or mental disability, we want to know more about the type of disability at Q.4. Some disabilities you will be able to observe readily. Other, like a learning disability or mental retardation, may be evident only if someone mentions the problem. It is also possible that a respondent could have more than one of the specified disabilities, so code all that apply. The following examples and definitions of some specific disabilities are to be used as guidelines to help you determine more accurately, the nature of the respondent's disability.

1 Learning disability: If a respondent says that he/she cannot participate in the exercise because of a learning disorder or learning disability, this code would apply. The most common type of learning disability is dyslexia, a type of impairment that affects a person's ability to read. If the respondent mentions dyslexia, or says that he/she has some other form of learning disability, this code would apply. Note that a learning disability should not prevent the respondent from participating in the Background Questionnaire interview. In addition, not all learning disorders would prevent a person from completing the exercise. For example, a person with Attention Deficit Disorder (ADD) is capable of reading and may have no difficulty with completing the Exercise. Persons reporting a learning disability should still be encouraged to attempt the tasks in the Exercise Booklet.

2 Any mental or emotional condition: This includes psychological disorders, dementia, autism, nervous breakdowns, severe depression, schizophrenia, etc.

3 Mental retardation: A person who is mentally slow or delayed in mental development. This can be caused by a birth defect, a congenital condition such as Down's Syndrome, or accident or injury to the brain that results in brain damage.

4 Hearing impairment: For R's who are totally (profoundly) deaf, this code should be used to describe a BQ nonresponse, since a totally deaf respondent cannot participate in the Background Questionnaire interview. R's who are partially deaf or hard of hearing, must be offered the opportunity to try the Exercise. If, however, they cannot begin or complete the Exercise because of a hearing impairment, use this code to describe the situation. Respondents with hearing impairments should be encouraged to use a hearing aid(s) or other device that would enable them to take part in the Background Questionnaire and Exercise.

5 Blind/Visual impairment: If R is totally blind, he/she can participate in the Background Questionnaire, but not in the Exercise, so this code would be used to describe an Exercise nonresponse. R's who have a visual impairment must be offered the opportunity to try the Exercise. If, however, they cannot begin or complete the Exercise because of a visual impairment, use this code to describe the situation. Respondents with visual impairments should be encouraged to use glasses or other devices that would enable them to take part in the Exercise.

6 Speech/Language impairment: Includes conditions such as severe stuttering that disable speech communication in English other than lack of proficiency in English. This code also includes aphasia (acquired impairment of verbal behavior caused by brain damage that could impair the R's ability to use and/or understand words).

7 A physical disability: Includes physical conditions that interfere with an ability to perform the Exercise such as paralysis, amputation of hand/arm, lack of muscle control, etc.

8 Other: Includes Alzheimer's disease, senility and any other condition not covered elsewhere in this code structure.

For a respondent who has had a stroke or suffers from another medical condition that would prevent him/her from completing the Exercise, you should code the symptom(s) the respondent suffers from. For example, a stroke victim who is left paralyzed, blind, and unable to speak would be coded as "5 - Blind/Visual impairment," "6 - Speech/Language impairment," and "7 - A physical disability."

Q.5 Respondents with reading and writing difficulties are encouraged to complete the Background Questionnaire and attempt the Exercise. Respondents who cannot read at all, or very little, may decide not to attempt the Exercise when you hand them the Exercise Booklet. If the respondent volunteers information on the extent of his or her reading ability at this time, record it here.

Q.6 If the respondent was unwilling to begin or to complete the Background Questionnaire or Exercise, record the reason for the refusal or the breakoff, using any verbatim response provided by the respondent or someone refusing for him or her. In the case of a breakoff, record where the breakoff occurred and the words used by the respondent to break off. If a language difficulty is behind the breakoff, note the language involved and describe the situation. Be sure to put a note in the Background Questionnaire or in the appropriate place in the Interview Guide at the point of the breakoff.

Q.7 Use this space to describe the circumstances behind a respondent being "unavailable for the field period," unable to contact after "maximum calls" or unable to be interviewed for "other reasons."

Q.8 You will answer this question for Background Questionnaire nonresponse situations, such as refusals, that might be "converted" if reassigned to another interviewer. This could be a recommendation for a reassignment interviewer of certain characteristics ("R might talk to older interviewer") or a conversation approach that the respondent might consider. As denoted by the skip patterns, language problems, disabilities, and reading/writing problems are skipped past this item because typically, they will not be reassigned.

Appendix F

FORM FOR INTERVIEWER'S OBERSERVATION ON THE EXERCISE BOOKLET AND THE OBSERVATIONS

Appendix F: Form for Interviewer's Observation on the Exercise Booklet and the Observations
NATIONAL ADULT LITERACY SURVEY

OBSERVATIONS ON EXERCISE BOOKLET

1. Did the respondent appear to try his or her best to complete the exercise booklet?

Yes	1
No	2

2. Did the respondent have difficulty ...

	Yes	No	N/A
using the newspaper?	1	2	3
using the almanac?	1	2	3
using the calculator (other than lighting)?	1	2	3
responding orally to the taped question(s)?	1	2	3

3. Did the respondent have any noticeable physical problems that could have affected his or her performance?

Yes	1
No	2

4. Were there any problems or disruptions during the administration of the exercise booklet that might have affected the respondent's performance?

Yes (briefly explain)	1
No	2

If you find this
Return it postage guaranteed to:

Westat
1650 Research Blvd.
Rockville, MD
20850-3129

6.11 Interviewer Observations

At the conclusion of the exercise, (after you have left the household), complete the observation page in the inside of the Call Record Folder. This page is shown as Exhibit 6-7.) Base your answers on the respondent's overall attitude and approach towards the exercise. Note that you only complete this final set of observations if the respondent completed the entire exercise (exercise result code "1"). In multi-respondent households, complete the Call Record observations for Respondent #1.

Exhibit 6-7 Interviewer Observations Section of Call Record Folder

INTERVIEWER OBSERVATIONS

If the Exercise Code is a "1", please answer the following questions. If more than one R in HH complete for R #1.

	Yes	No
1. Were you able to conduct the exercise in a place with adequate lighting?	1	2
2. Were there many disruptions during the administration of the exercises that might have effected the R's performance?	1	2
3. Did the R seem interested in participating in the survey?	1	2
4. Did the R appear to try his or her best to complete the set of exercises?	1	2
5. Did the R have any noticeable physical problems that could have affected his or her performance?	1	2

NOTES

Appendix G

ENGLISH BACKGROUND QUESTIONNAIRE FOR HOUSEHOLDS

USING PRINTED MATERIALS

A SURVEY OF AMERICA'S ADULTS

NALS

PLACE LABEL HERE

⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	
①	①	①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②	②	②
③	③	③	③	③	③	③	③	③	③	
④	④	④	④	④	④	④	④	④	④	
⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	
⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	
⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	
⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	
⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	

ENGLISH BACKGROUND QUESTIONNAIRE

Booklet ID #					
⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①
②	②	②	②	②	②
③	③	③	③	③	③
④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨

Office Use Only					
⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①
②	②	②	②	②	②
③	③	③	③	③	③
④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨

BQ Elapsed Time		
⓪	⓪	⓪
①	①	①
②	②	②
③	③	③
④	④	④
⑤	⑤	⑤
⑥	⑥	⑥
⑦	⑦	⑦
⑧	⑧	⑧
⑨	⑨	⑨

EX Elapsed Time		
⓪	⓪	⓪
①	①	①
②	②	②
③	③	③
④	④	④
⑤	⑤	⑤
⑥	⑥	⑥
⑦	⑦	⑦
⑧	⑧	⑧
⑨	⑨	⑨

The National Center for Education Statistics is supporting the conduct of NALS through a contract awarded to

Educational Testing Service

Copyright © 1991 by Educational Testing Service. All Rights Reserved.

OMB No. 1859-0654. Approval Expires August 31, 1993

A Survey Of America's Adults
BACKGROUND QUESTIONNAIRE

IF R IS NOT SCREEENER R:	IF R IS SCREEENER R:
Hello, I am (NAME) from Westat. My organization is helping the United States Department of Education with a very important survey about how adults use printed materials. Recently, another member of your household told me who lives here. Based on this information, you were selected at random to take part in the survey.	You have been selected to participate in the survey.

I will ask you a short set of questions about your background, education, and work experiences. Then, I will give you a booklet containing exercises based on printed materials, such as newspapers, maps, stories, brochures, forms, and advertisements. Others who have completed these exercises found them interesting and fun. The entire survey will take approximately one hour to complete, and you will be paid $20 for your participation.

Your participation in this survey is voluntary and very important. All of your answers will be kept strictly confidential. All information will be reported for a group as a whole and your responses will not be linked to your name.

IF R REQUESTS MORE INFORMATION ABOUT PURPOSE OF SURVEY:
Today, adults increasingly are expected to use printed information in our society, but there is very little information available on whether or not they are well prepared. This survey will provide information about the reading and writing experiences, activities, and skills of adults in the United States. Information will be used by educators, policymakers, and business leaders to design programs in order to improve the literacy skills of adults.

INITIAL
INTERVIEWER NAME: ______________________________

REASSIGNMENT
INTERVIEWER NAME: ______________________________

Scorer #1		
⓪	⓪	⓪
①	①	①
②	②	②
③	③	③
④	④	④
⑤	⑤	⑤
⑥	⑥	⑥
⑦	⑦	⑦
⑧	⑧	⑧
⑨	⑨	⑨

Scorer #2		
⓪	⓪	⓪
①	①	①
②	②	②
③	③	③
④	④	④
⑤	⑤	⑤
⑥	⑥	⑥
⑦	⑦	⑦
⑧	⑧	⑧
⑨	⑨	⑨

START TIME: ______ A.M. ______ P.M.	

SECTION A
GENERAL AND LANGUAGE BACKGROUND

A-1. In what country were you born?

USA (50 STATES OR D.C.)	①	(A-4)
U.S. TERRITORY (SPECIFY): __________	②	(A-2)
OTHER (SPECIFY COUNTRY): __________	③	(A-2)

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

A-2. How many years have you lived in the United States?

1 TO 5	①
6 TO 10	②
11 TO 15	③
16 TO 20	④
21 TO 30	⑤
31 TO 40	⑥
41 TO 50	⑦
51 OR MORE	⑧

A-3. What was the highest level of education you completed before coming to the United States?
(IF RESPONSE DOES NOT FIT CATEGORIES, PROBE FOR EQUIVALENT.)

DID NOT ATTEND SCHOOL BEFORE COMING TO U.S.	①
PRIMARY (GRADES K-3)	②
ELEMENTARY (GRADES 4-8)	③
SECONDARY (GRADES 9-12)	④
VOCATIONAL (POSTSECONDARY)	⑤
COLLEGE/UNIVERSITY	⑥
OTHER (SPECIFY): ________________	⑦

A-4. When you were growing up, what language or languages were usually spoken in your home?
(CODE ALL THAT APPLY.)

ENGLISH	①
SPANISH	②
OTHER (SPECIFY): ________________	③

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

A-5. What language or languages did you learn to speak before you started school? (CODE ALL THAT APPLY.)

ENGLISH .... ①
SPANISH .... ②
OTHER (SPECIFY): __________ ③

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

BOX 1. IF ENGLISH ONLY FOR BOTH A-4 AND A-5, SKIP TO A-17; OTHERWISE, CONTINUE.

A-6. What language did you first learn to read and write?

ENGLISH .... ①
SPANISH .... ②
OTHER (SPECIFY): __________ ③
NEVER LEARNED TO READ AND WRITE .. ④

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

A-7. How old were you when you learned to speak English?

1-4 YEARS .... ①
5-10 YEARS .... ②
11-15 YEARS .... ③
16-20 YEARS .... ④
21 OR OLDER .... ⑤
DOES NOT SPEAK ENGLISH .... ⑥

BOX 2. IF ENGLISH ONLY IN A-5, SKIP TO A-17; OTHERWISE, CONTINUE. IF MORE THAN ONE NON-ENGLISH LANGUAGE IN A-5, USE FIRST ONE MENTIONED FOR A-8, A-9, AND A-10.

A-8. With regard to (NON-ENGLISH LANGUAGE IN A-5), how well do you…

Would you say…

	very well	well	not well	or	not at all
understand it when it is spoken to you?	①	②	③		④
speak it?	①	②	③		④
read it?	①	②	③		④
write it?	①	②	③		④

A-9. HAND CARD: With regard to (NON-ENGLISH LANGUAGE IN A-5), how often do you…

HAND CARD 1

Would you say…

	every day	a few times a week	once a week	less than once a week	or never
A. listen to radio programs, tapes, or records in (LANGUAGE)?	①	②	③	④	⑤
B. watch television programs or video tapes in (LANGUAGE)?	①	②	③	④	⑤
C. read newspapers, magazines, or books in (LANGUAGE)?	①	②	③	④	⑤
D. write or fill out letters or forms in (LANGUAGE)?	①	②	③	④	⑤

A-10. HAND CARD: Looking at the categories on this card, tell me what language you use in each of the following situations.

HAND CARD 2

Would you say…

	always English	more English than (LANGUAGE)	English and (LANGUAGE) equally	more (LANGUAGE) than English	or always (LANGUAGE)	DOES NOT APPLY
A. At home	①	②	③	④	⑤	⑥
B. At work	①	②	③	④	⑤	⑥
C. While shopping in your neighborhood	①	②	③	④	⑤	⑥
D. When visiting relatives or friends	①	②	③	④	⑤	⑥

A-11. Have you ever taken a course to learn how to read and write English as a second language?

YES .... ① (A-12)
NO .... ② (A-13)

A-12. Did you complete this course?

YES .... ①
NO .... ②

A-13. Have you ever taken a course to learn how to speak and understand English as a second language?

YES .... ① (A-14)
NO .... ② (A-15)

A-14. Did you complete this course?

YES .... ①
NO .... ②

A-15. Which language do you usually speak now?

ENGLISH .... ①
SPANISH .... ②
OTHER (SPECIFY): ________________ ③

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

A-16. What other language do you often speak now?

ENGLISH .... ①
SPANISH .... ②
OTHER (SPECIFY): ________________ ③
NO OTHER LANGUAGE SPOKEN .... ④

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

A-17. With regard to the English language, how well do you…

Would you say…

	very well	well	not well	or	not at all
understand it when it is spoken to you?	①	②	③		④
speak it?	①	②	③		④
read it?	①	②	③		④
write it?	①	②	③		④
do arithmetic problems when you have to get the numbers from written materials?	①	②	③		④

SECTION B
EDUCATIONAL BACKGROUND AND EXPERIENCES

B-1. HAND CARD: I'd like to ask you about your educational background and experiences. What is the highest level of public or private education you completed? (IF R WENT TO SCHOOL OUTSIDE U.S., PROBE FOR EQUIVALENT.)

HAND CARD 3

A.	STILL IN HIGH SCHOOL	(01)	(B-4)
B.	LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE): ⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧	(02)	(B-2)
C.	SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT COMPLETE 12TH GRADE) (SPECIFY GRADE): ⑨ ⑩ ⑪	(03)	(B-2)
D.	GED OR HIGH SCHOOL EQUIVALENCY	(04)	(B-2)
E.	HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM)	(05)	(B-4)
F.	ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL	(06)	
G.	COLLEGE: LESS THAN TWO YEARS	(07)	
H.	COLLEGE: ASSOCIATE'S DEGREE (A.A.)	(08)	
I.	COLLEGE: TWO YEARS OR MORE, NO DEGREE	(09)	
J.	COLLEGE GRADUATE (B.S. OR B.A.)	(10)	
K.	POSTGRADUATE/NO DEGREE	(11)	
L.	POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.)	(12)	
M.	DON'T KNOW	(13)	(B-5)

B-2. What was the main reason you stopped your public or private schooling when you did?

FINANCIAL PROBLEMS	①
WENT TO WORK OR INTO THE MILITARY	②
PREGNANCY	③
LOST INTEREST OR BEHAVIOR PROBLEMS IN SCHOOL	④
ACADEMIC PROBLEMS AT SCHOOL	⑤
FAMILY OR PERSONAL PROBLEMS	⑥
OTHER	⑦

BOX 3. CHECK B-1. IF CODED 04, SKIP TO B-5; OTHERWISE, CONTINUE.

B-3. Did you ever study for a GED or high school equivalency?

YES	①
NO	②

BOX 4. CHECK B-1. IF CODED 02, SKIP TO B-5; OTHERWISE, CONTINUE.

B-4. What kind of high school program (are you taking/did you take)? (Is/Was) it…

college preparatory,	①
vocational, technical, or trade,	②
general,	③
or the same for everyone?	④

BOX 5. CHECK B-1. IF CODED 01, SKIP TO B-6; OTHERWISE, CONTINUE.

B-5. Are you currently enrolled in school or college, either full-time or part-time?

YES	①	(B-6)
NO	②	(B-7)

B-6. Upon completing your current schooling, what diplomas, certificates, degrees, or accreditation do you expect to earn?

HIGH SCHOOL DIPLOMA OR EQUIVALENCY	①
VOCATIONAL, TRADE, OR BUSINESS	②
TWO YEARS OF COLLEGE (ASSOCIATE'S DEGREE)	③
FOUR- OR FIVE-YEAR COLLEGE DEGREE (B.S., B.A.)	④
MASTER'S, PH.D., M.D., OR OTHER ADVANCED DEGREE	⑤
OTHER	⑥
NONE	⑦

B-7. Are you currently enrolled in or have you ever taken part in a program other than in regular school in order to improve your *basic skills*, that is, basic reading, writing, and arithmetic skills?

YES	①	(B-8)
NO	②	(B-10)

ASK B-8 AND B-9 IN SEQUENCE FOR EACH PROGRAM CODED "YES" IN B-8.

B-8. (Is/Was) this program…	YES	NO	DON'T KNOW	**B-9.** IF YES: How recently did you take part in this training?
a training program or courses given or sponsored by your employer or union?	①	②	③	STILL ENROLLED ① WITHIN THE PAST YEAR ② BETWEEN ONE & FIVE YEARS ③ MORE THAN FIVE YEARS AGO ④
a publicly sponsored education and training program, such as JTPA or ABE?	①	②	③	STILL ENROLLED ① WITHIN THE PAST YEAR ② BETWEEN ONE & FIVE YEARS ③ MORE THAN FIVE YEARS AGO ④
a tutoring program sponsored by a library, church, or community organization?	①	②	③	STILL ENROLLED ① WITHIN THE PAST YEAR ② BETWEEN ONE & FIVE YEARS ③ MORE THAN FIVE YEARS AGO ④
any other program (such as one offered by the military, prisons, or other institutions)?	①	②	③	STILL ENROLLED ① WITHIN THE PAST YEAR ② BETWEEN ONE & FIVE YEARS ③ MORE THAN FIVE YEARS AGO ④

B-10. HAND CARD: Where did you primarily learn to…

HAND CARD
4

Would you say…

	mostly in school	mostly at home or in the community	mostly at work	or	did not learn	OTHER
A. read newspapers, magazines, or books?	①	②	③		④	⑤
B. read graphs, diagrams, or maps?	①	②	③		④	⑤
C. fill out forms?	①	②	③		④	⑤
D. write letters, notes, memos, or reports?	①	②	③		④	⑤

B-11. Do you have any health problem, impairment, or disability now that keeps you from participating fully in work, school, housework, or other activities?

YES .... ① (B-12)
NO .... ② (C-1)

B-12. Do you have any difficulty seeing the words and letters in ordinary newspaper print even when wearing glasses or contact lenses, if you usually wear them?

YES .... ①
NO .... ②

B-13. Do you have any difficulty hearing what is said in a normal conversation with another person even when using a hearing aid, if you usually wear one?

YES .... ①
NO .... ②

B-14. Do you currently have any of these conditions? (READ LIST AND CODE YES OR NO FOR EACH CONDITION.)

	YES	NO
A learning disability	①	②
An emotional disorder	①	②
Mental retardation	①	②
A speech disability	①	②
A physical disability	①	②
A long-term illness (6 months or more)	①	②
Other health impairment	①	②

SECTION C
POLITICAL AND SOCIAL PARTICIPATION

C-1. I'd like to find out how you usually get information about current events, public affairs, and the government. HAND CARD: How much information do you get from… (REPEAT QUESTION AS NECESSARY.)

HAND CARD 5

Would you say…

	a lot	some	a little	or	none
A. newspapers?	①	②	③		④
B. magazines?	①	②	③		④
C. radio?	①	②	③		④
D. television?	①	②	③		④
E. family members, friends, or co-workers?	①	②	③		④

C-2. How many hours do you usually watch television each day?

NONE	①
1 HOUR OR LESS	②
2 HOURS	③
3 HOURS	④
4 HOURS	⑤
5 HOURS	⑥
6 OR MORE HOURS	⑦

C-3. How often do you use the services of a library, for any reason? Would you say daily, weekly, monthly, once or twice a year, or never?

DAILY	①
WEEKLY	②
MONTHLY	③
ONCE OR TWICE A YEAR	④
NEVER	⑤

C-4. Have you voted in a national or state election in the United States within the past five years?

YES	①
NO	②
NOT ELIGIBLE TO VOTE	③

SECTION D
LABOR FORCE PARTICIPATION

D-1. Now I'd like to ask you some questions about what you were doing last week. HAND CARD: Last week were you... (CODE ALL THAT APPLY AND FOLLOW SKIP INSTRUCTION FOR LOWEST OPTION CODED.)

HAND CARD 6

A.	working a full-time job for pay or profit, that is, 35 hours or more?	01	(D-3)
B.	working for pay or profit part-time, that is, 1 to 34 hours?	02	(D-2)
C.	working two or more part-time jobs for pay, totaling 35 or more hours?	03	(D-2)
D.	unemployed, laid off, or looking for work?	04	(D-2)
E.	with a job but not at work because of temporary illness, vacation, or work stoppage?	05	(D-3)
F.	with a job but on family leave (maternity or paternity leave)?	06	(D-3)
G.	in school?	07	(D-5)
H.	keeping house?	08	(D-5)
I.	retired?	09	(D-5)
J.	doing volunteer work?	10	(D-5)
K.	OTHER (SPECIFY): ____________	11	(D-5)

D-2. Have you looked for a job at any time during the past four weeks?

YES	①
NO	②

BOX 6. CHECK D-1. IF CODED 04, SKIP TO D-5; OTHERWISE, CONTINUE.

D-3. Last week, what was your total weekly wage or salary from all jobs *before* any deductions? Include tips and commissions. (WRITE IN DOLLAR AMOUNT AND CHECK APPROPRIATE CODE.)

$__________ . _________

PER HOUR	①	(D-4)
PER DAY	②	(D-4)
PER WEEK	③	(D-5)
PER TWO-WEEK PERIOD	④	(D-5)
PER MONTH	⑤	(D-5)
PER YEAR	⑥	(D-5)
OTHER (SPECIFY):____________	⑦	(D-5)

⓪	⓪	⓪	⓪	⓪
①	①	①	①	①
②	②	②	②	②
③	③	③	③	③
④	④	④	④	④
⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨

D-4. How many (hours/days) did you work last week?

SPECIFY: ______________________

D-5. Now I'd like to ask you some questions about your work during the past 12 months. Including weeks of paid leave, such as vacation and sick leave, how many weeks did you work for pay or profit during the past 12 months?

NONE (0)	①	(D-6)
LESS THAN 52 WEEKS (SPECIFY NUMBER OF WEEKS): __________	②	(D-6)
52 WEEKS (FOR THE LAST 12 MONTHS)	③	(D-7)

⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨
⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨

D-6. Of the weeks you were not employed, what were you doing? Were you…
(CODE ALL THAT APPLY.)

looking for work,	①
in school or training,	②
keeping house,	③
retired,	④
volunteering,	⑤
or doing something else? (SPECIFY): __________	⑥

BOX 7. CHECK D-5. IF CODED 1, GO TO D-9; OTHERWISE, CONTINUE.

D-7. For the past 12 months, what was your average weekly wage or salary *before* any deductions? Include tips and commissions. (WRITE IN DOLLAR AMOUNT AND CHECK APPROPRIATE CODE.)

$__________ . __________

PER HOUR	①	(D-8)
PER DAY	②	(D-8)
PER WEEK	③	(D-10)
PER TWO-WEEK PERIOD	④	(D-10)
PER MONTH	⑤	(D-10)
PER YEAR	⑥	(D-10)
OTHER (SPECIFY): __________	⑦	(D-10)

⓪	⓪	⓪	⓪	⓪
①	①	①	①	①
②	②	②	②	②
③	③	③	③	③
④	④	④	④	④
⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨

D-8. On average, how many (hours/days) did you work during the past 12 months?

SPECIFY: ____________________ (D-10)

D-9. Which of the following describes your work history? Have you…

held a paying job within the last three years?	①	(D-10)
held a paying job, but not within the last three years?	②	(E-1)
never been employed for pay full-time or part-time?	③	

D-10. Now I'd like to ask you some questions about your current full-time or part-time job or your most recent full-time or part-time job. For what kind of business or industry (do/did) you work? (For example, television and radio manufacturing, retail shoe store, state labor department, farm, etc.) (IF R IS WORKING TWO OR MORE JOBS, PROBE: Tell me about the job you work the most hours or the job you consider your primary employment.)

BUSINESS OR INDUSTRY: ____________________

____

____

____

⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨
⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨
⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨

D-11. What (is/was) your occupation, that is, what (is/was) your job called? (For example, electrical engineer, stock clerk, typist, farmer, etc.)

WORK: ____________________________________

____

____

____

⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨
⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨
⓪	①	②	③	④	⑤	⑥	⑦	⑧	⑨

D-12. What (are/were) the most important activities or duties at this job? (For example, typing, keeping account books, filing, selling cars, operating a printing press, finishing concrete, etc.)

ACTIVITIES OR DUTIES: ____________________

____

____

____

____

____

SECTION E
LITERARY ACTIVITIES AND COLLABORATION

E-1. Now I'd like to talk to you about what you read in English. First, let's talk about newspapers. How often do you read a newspaper — every day, a few times a week, once a week, less than once a week, or never?

EVERY DAY	①	(E-2)
A FEW TIMES A WEEK	②	(E-2)
ONCE A WEEK	③	(E-2)
LESS THAN ONCE A WEEK	④	(E-2)
NEVER	⑤	(E-3)

E-2. Now I will read a list of the different parts of newspapers. Would you please tell me which parts you generally read when looking at a newspaper? Do you read... (CODE ALL THAT APPLY.)

news?	01
sports?	02
home, fashion, or health?	03
editorial pages?	04
financial news or stock listings?	05
comics?	06
classified ads or other advertisements?	07
TV, movie, or concert listings?	08
book, movie, or art reviews?	09
horoscope or advice columns?	10
something else? (SPECIFY): ________________	11
NONE	12

E-3. Next, let's talk about magazines. About how many different magazines do you look at or read in English on a regular basis?

0	①
1	②
2	③
3 TO 5	④
6 OR MORE	⑤

E-4. Next, I'd like to ask you about books you may have read in English recently. I will read a list of different types of books. Would you please tell me which of these types of books you've read in English in the past six months, if any? Have you read... (CODE ALL THAT APPLY.) (INCLUDE COURSE BOOKS.)

fiction?	①
recreation or entertainment?	②
current affairs or history?	③
inspiration or religion?	④
science or social science?	⑤
reference, such as encyclopedias or dictionaries?	⑥
manuals for cooking, operating, repairing, or building?	⑦
any other types of books? (SPECIFY): ________________	⑧
NONE	⑨

I would like to ask you about other kinds of things you may have read in English, first for your own use and then for work, if appropriate. Please consider printed or written materials, as well as information displayed on computer screens.

E-5. HAND CARD: How often do you read or use information from (ITEM) for your own use? Would you say every day, a few times a week, once a week, less than once a week, or never?

HAND CARD 1

	EVERY DAY	A FEW TIMES A WEEK	ONCE A WEEK	LESS THAN ONCE A WEEK	NEVER
A. Letters or memos	①	②	③	④	⑤
B. Reports, articles, magazines, or journals	①	②	③	④	⑤
C. Manuals or reference books, including catalogs or parts lists	①	②	③	④	⑤
D. Directions or instructions for medicines, recipes, or other products	①	②	③	④	⑤
E. Diagrams or schematics	①	②	③	④	⑤
F. Bills, invoices, spreadsheets, or budget tables	①	②	③	④	⑤

BOX 8. CHECK D-9 ON PAGE 13.

D-9 IS BLANK OR CODED 1. .... ① **(E-6)**

D-9 IS CODED 2 OR 3. .... ② **(E-7)**

E-6. HAND CARD: How often (do/did) you read or use information from (ITEM) as part of your (current/most recent) job? Would you say every day, a few times a week, once a week, less than once a week, or never?

HAND CARD 1

	EVERY DAY	A FEW TIMES A WEEK	ONCE A WEEK	LESS THAN ONCE A WEEK	NEVER
A. Letters or memos	①	②	③	④	⑤
B. Reports, articles, magazines, or journals	①	②	③	④	⑤
C. Manuals or reference books, including catalogs or parts lists	①	②	③	④	⑤
D. Directions or instructions for medicines, recipes, or other products	①	②	③	④	⑤
E. Diagrams or schematics	①	②	③	④	⑤
F. Bills, invoices, spreadsheets, or budget tables	①	②	③	④	⑤

Now I would like to find out about things that you write or fill out in English, first for your own use and then for work, if appropriate. Please consider writing or typing on paper, as well as entering information into computers.

E-7. HAND CARD: How often do you write or fill out (ITEM) for your own use? Would you say every day, a few times a week, once a week, less than once a week, or never?

HAND CARD 1

	EVERY DAY	A FEW TIMES A WEEK	ONCE A WEEK	LESS THAN ONCE A WEEK	NEVER
A. Letters or memos	①	②	③	④	⑤
B. Forms or things such as bills, invoices, or budgets	①	②	③	④	⑤
C. Reports or articles	①	②	③	④	⑤

BOX 9. CHECK BOX 8 ON PREVIOUS PAGE. IF CODED 2, SKIP TO E-9; IF CODED 1, CONTINUE WITH E-8.

E-8. HAND CARD: How often (do/did) you write or fill out (ITEM) as part of your (current/most recent) job? Would you say every day, a few times a week, once a week, less than once a week, or never?

HAND CARD 1

	EVERY DAY	A FEW TIMES A WEEK	ONCE A WEEK	LESS THAN ONCE A WEEK	NEVER
A. Letters or memos	①	②	③	④	⑤
B. Forms or things such as bills, invoices, or budgets	①	②	③	④	⑤
C. Reports or articles	①	②	③	④	⑤

E-9. How often do you use arithmetic or mathematics for your own use, that is, add, subtract, multiply, divide, or measure? Would say every day, a few times a week, once a week, less than once a week, or never?

EVERY DAY	①
A FEW TIMES A WEEK	②
ONCE A WEEK	③
LESS THAN ONCE A WEEK	④
NEVER	⑤

BOX 10. CHECK BOX 8 ON PAGE 16. IF CODED 2, SKIP TO E-11; IF CODED 1, CONTINUE WITH E-10.

E-10. How often (do/did) you use arithmetic or mathematics as part of your (current/most recent) job? Would say every day, a few times a week, once a week, less than once a week, or never?

EVERY DAY .... ①
A FEW TIMES A WEEK .... ②
ONCE A WEEK .... ③
LESS THAN ONCE A WEEK .... ④
NEVER .... ⑤

E-11. HAND CARD: How much help do you get from family members or friends with…

HAND CARD 5

Would you say…

	a lot	some	a little	or	none
A. filling out forms?	①	②	③		④
B. reading or explaining newspaper articles or other written information?	①	②	③		④
C. printed information associated with government agencies, public companies, private businesses, hospitals, etc.?	①	②	③		④
D. writing notes and letters?	①	②	③		④
E. using basic arithmetic, that is, adding, subtracting, multiplying, or dividing, such as in filling out order forms or balancing a checkbook?	①	②	③		④

SECTION F
DEMOGRAPHIC INFORMATION

F-1. What was the highest level of education your mother (stepmother or female guardian) completed? (IF WENT TO SCHOOL OUTSIDE U.S., PROBE FOR EQUIVALENT.)

LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE):	01
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧	
SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT COMPLETE 12TH GRADE) (SPECIFY GRADE):	02
⑨ ⑩ ⑪	
GED OR HIGH SCHOOL EQUIVALENCY	03
HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM)	04
ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL	05
COLLEGE: LESS THAN TWO YEARS	06
COLLEGE: ASSOCIATE'S DEGREE (A.A.)	07
COLLEGE: TWO YEARS OR MORE, NO DEGREE	08
COLLEGE GRADUATE (B.S. OR B.A.)	09
POSTGRADUATE/NO DEGREE	10
POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.)	11
DON'T KNOW	12

F-2. What was the highest level of education your father (stepfather or male guardian) completed? (IF WENT TO SCHOOL OUTSIDE U.S., PROBE FOR EQUIVALENT.)

LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE):	01
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧	
SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT COMPLETE 12TH GRADE) (SPECIFY GRADE):	02
⑨ ⑩ ⑪	
GED OR HIGH SCHOOL EQUIVALENCY	03
HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM)	04
ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL	05
COLLEGE: LESS THAN TWO YEARS	06
COLLEGE: ASSOCIATE'S DEGREE (A.A.)	07
COLLEGE: TWO YEARS OR MORE, NO DEGREE	08
COLLEGE GRADUATE (B.S. OR B.A.)	09
POSTGRADUATE/NO DEGREE	10
POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.)	11
DON'T KNOW	12

I would now like to ask you some questions about your household.

F-3. HAND CARD: First, which letter on this card describes your current marital status?

HAND CARD 7

A. NEVER MARRIED .... ①
B. MARRIED, LIVING WITH SPOUSE .... ②
C. MARRIED, SPOUSE LIVING ELSEWHERE .... ③
D. LIVING AS MARRIED .... ④
E. SEPARATED OR DIVORCED .... ⑤
F. WIDOWED .... ⑥

F-4. Including yourself, how many people in your household are employed or work for pay or wages part-time (1 to 34 hours)?

NONE .... ①
ONE .... ②
TWO .... ③
THREE OR MORE .... ④

F-5. Including yourself, how many people in your household are employed or work for pay or wages full-time (35 hours or more)?

NONE .... ①
ONE .... ②
TWO .... ③
THREE OR MORE .... ④

F-6. HAND CARD: Did you or anyone in your household receive any of the following during the past 12 months? (CODE ALL THAT APPLY.)

HAND CARD 8

A. Social Security or Railroad Retirement payments .... ①
B. Supplemental Security Income (SSI) .... ②
C. Other retirement, survivor, or disability payments (other than Social Security or Railroad Retirement) .... ③
D. Aid to Families with Dependent Children (AFDC), public assistance, or public welfare payments from the state or local welfare office .... ④
E. Food stamps .... ⑤
F. Interest from savings or other bank accounts (other than dividends) .... ⑥
G. Dividend income from stocks or mutual funds or income from rental property, royalty, estates, or trusts .... ⑦
H. Income from any other sources, such as Veterans Administration payments, workers or unemployment compensation, child support, or alimony .... ⑧
I. NONE .... ⑨

F-7. HAND CARD: Which letter on this card corresponds to your approximate personal income from all sources for the past 12 months?

HAND CARD 9

A. UNDER $5,000	01
B. $ 5,000 - $ 9,999	02
C. $10,000 - $14,999	03
D. $15,000 - $19,999	04
E. $20,000 - $29,999	05
F. $30,000 - $39,999	06
G. $40,000 - $49,999	07
H. $50,000 - $74,999	08
I. $75,000 - $99,999	09
J. $100,000 AND OVER	10
K. REFUSED	11
L. DON'T KNOW	12
M. NO PERSONAL INCOME	13

F-8. HAND CARD: Which letter on this card corresponds to your approximate total household income from all sources for the past 12 months? (IF UNDERGRADUATE COLLEGE STUDENT LIVING AWAY FROM "HOME": Please provide household income for your permanent residence.)

HAND CARD 9

A. UNDER $5,000	01
B. $ 5,000 - $ 9,999	02
C. $10,000 - $14,999	03
D. $15,000 - $19,999	04
E. $20,000 - $29,999	05
F. $30,000 - $39,999	06
G. $40,000 - $49,999	07
H. $50,000 - $74,999	08
I. $75,000 - $99,999	09
J. $100,000 AND OVER	10
K. REFUSED	11
L. DON'T KNOW	12

F-9 HAND CARD: Which of the groups on this card best describes you? (IF RESPONDENT REFUSES TO ANSWER, PLEASE RECORD THE RESPONDENT'S RACE FROM OBSERVATION.)

HAND CARD 10

A. WHITE	1
B. BLACK, AFRICAN AMERICAN	2
C. AMERICAN INDIAN	3
D. ALASKAN NATIVE	4
E. PACIFIC ISLANDER	5
F. ASIAN (SPECIFY): ______	6
G. OTHER (SPECIFY): ______	7

F-10. Are you of Spanish or Hispanic origin or descent?

YES .... ① (F-11)
NO .... ② (F-12)

F-11. HAND CARD: Which of the groups on this card best describes your Hispanic origin?

HAND CARD 11

A. MEXICANO, MEXICAN, MEXICAN AMERICAN, CHICANO .... ①
B. PUERTO RICAN .... ②
C. CUBAN .... ③
D. CENTRAL/SOUTH AMERICAN .... ④
E. OTHER SPANISH/HISPANIC (SPECIFY): ____________ ⑤

F-12. What is your date of birth?

____
MONTH / DAY / YEAR

Date of Birth					
⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①
	②	②	②	②	②
	③	③	③	③	③
	④		④	④	④
	⑤		⑤	⑤	⑤
	⑥		⑥	⑥	⑥
	⑦		⑦	⑦	⑦
	⑧		⑧	⑧	⑧
	⑨		⑨	⑨	⑨

INTERVIEWER PLEASE NOTE:

F-13. SEX

MALE .... ①
FEMALE .... ②

END
TIME: ______ A.M. ______ P.M.

BOX 11. IS THIS INTERVIEW BEING CONDUCTED IN:

CALIFORNIA
ILLINOIS
INDIANA
IOWA
LOUISIANA
NEW JERSEY
NEW YORK
OHIO
PENNSYLVANIA
TEXAS
WASHINGTON

YES .... ① **(GO TO TABLE OF CONTENTS ON NEXT PAGE.)**
NO .... ② **(CONTINUE WITH EXERCISE.)**

BOX 12. LOCATE THE PAGE OF QUESTIONS APPROPRIATE TO THE STATE IN WHICH YOU ARE INTERVIEWING AND SAY: Now I have just a few questions to ask that are especially important to your state; THEN CONTINUE WITH STATE QUESTIONS.

CALIFORNIA	PAGE 24
ILLINOIS	PAGE 25
INDIANA	PAGE 26
IOWA	PAGE 27
LOUISIANA	PAGE 28
NEW JERSEY	PAGE 29
NEW YORK	PAGE 30
OHIO	PAGE 31
PENNSYLVANIA	PAGE 32
TEXAS	PAGE 33
WASHINGTON	PAGE 34

CALIFORNIA

CA-1. How many years have you lived in California?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

CA-2. How likely is it that you will move out of California in the next five years? Would you say…

not likely,	①
somewhat likely, or	②
very likely?	③

CA-3. Which of the following statements apply to you? (CODE ALL THAT APPLY.)

I grew up in a home where there were a variety of books, newspapers, magazines, and other printed materials.	①
I was read to as a young child by my parent(s) or guardian(s).	②
I was helped with schoolwork by my parent(s) or guardian(s).	③
NONE OF THE ABOVE	④

CA-4. Do you feel that an employer has an obligation to provide literacy education to its employees who need assistance?

YES	①
NO	②
NO OPINION	③

CA-5. If you have any children under six years of age living with you now, how often do you read to or with them?

EVERY DAY	①
A FEW TIMES A WEEK	②
ABOUT ONCE A WEEK	③
ABOUT ONCE OR TWICE A MONTH	④
ALMOST NEVER	⑤
RESPONDENT HAS NO CHILDREN UNDER SIX YEARS OF AGE IN THE HOUSEHOLD.	⑥

ILLINOIS

IL-1. Which of the following levels of schooling did you attend in Illinois? (CODE ALL THAT APPLY.)

Kindergarten through 3rd grade ①
Grades 4 through 8 ②
Grades 9 through 12 ③
Vocational school or community college (One- or two-year program) ④
Four-year college or university ⑤
Did not attend school in Illinois ⑥

IL-2. HAND CARD: Where would the best location be for you to take a course or a training program?

HAND CARD 12

A LOCAL PUBLIC SCHOOL OUTSIDE OF SCHOOL HOURS, ①
A LOCAL COMMUNITY COLLEGE, ②
A UNIVERSITY, ③
A LIBRARY, CHURCH, COMMUNITY CENTER, OR ADULT EDUCATION CENTER, ④
YOUR WORKPLACE, OR ⑤
SOME OTHER LOCATION? ⑥
NOT INTERESTED IN TAKING A COURSE ⑦

IL-3. For which of the following reasons would you enroll in a course or a training program? (CODE ALL THAT APPLY.)

To qualify for a better job ①
To participate in leisure activities or for personal improvement ②
To complete a diploma or degree, or to qualify for a license ③
To improve job-related skills ④
To improve parenting/child-care skills ⑤
NONE OF THE ABOVE ⑥

IL-4. HAND CARD: Which one of the following reasons would most likely keep you from enrolling in a course or a training program?

HAND CARD 13

LACK OF TIME OR INTEREST, ①
COURSE OR TRAINING PROGRAM NOT OFFERED AT CONVENIENT TIMES, ②
INABILITY TO PAY FOR COURSE OR TRAINING, ③
LACK OF CHILD CARE, ④
LACK OF TRANSPORTATION, OR ⑤
LACK OF INFORMATION ABOUT AVAILABLE COURSES OR TRAINING PROGRAMS? ⑥
NONE OF THE ABOVE ⑦

IL-5. If you are currently employed, in which of the following areas do you feel you need additional training to become a more productive worker in your present job? (CODE ALL THAT APPLY.)

Reading ①
Arithmetic ②
Computers ③
Problem solving ④
Listening ⑤
Working in teams ⑥
None of the above ⑦
NOT CURRENTLY EMPLOYED ⑧

INDIANA

IN-1. How many years have you lived in Indiana?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

IN-2. Which of the following levels of schooling did you attend in Indiana? (CODE ALL THAT APPLY.)

Kindergarten through 3rd grade	①
Grades 4 through 8	②
Grades 9 through 12	③
Vocational school or community college (One- or two-year program)	④
Four-year college or university	⑤
Did not attend school in Indiana	⑥

IN-3. Which of the following statements apply to you? (CODE ALL THAT APPLY.)

I grew up in a home where there were a variety of books, newspapers, magazines, and other printed materials.	①
I was read to as a young child by my parent(s) or guardian(s).	②
I was helped with schoolwork by my parent(s) or guardian(s).	③
NONE OF THE ABOVE	④

IN-4. If you have any children under six years of age living with you now, how often do you read to or with them?

EVERY DAY	①
A FEW TIMES A WEEK	②
ABOUT ONCE A WEEK	③
ABOUT ONCE OR TWICE A MONTH	④
ALMOST NEVER	⑤
RESPONDENT HAS NO CHILDREN UNDER SIX YEARS OF AGE IN THE HOUSEHOLD.	⑥

IN-5. HANDCARD: Which one of the following reasons would most likely keep you from enrolling in a course or a training program?

HAND CARD 13

LACK OF TIME OR INTEREST,	①
COURSE OR TRAINING PROGRAM NOT OFFERED AT CONVENIENT TIMES,	②
INABILITY TO PAY FOR COURSE OR TRAINING,	③
LACK OF CHILD CARE,	④
LACK OF TRANSPORTATION, OR	⑤
LACK OF INFORMATION ABOUT AVAILABLE COURSES OR TRAINING PROGRAMS?	⑥
NONE OF THE ABOVE	⑦

IOWA

IA-1. How many years have you lived in Iowa?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

IA-2. HAND CARD: Which one of the following would be the most important reason for you not to take part in a basic skills program?

HAND CARD 14

I DON'T THINK I NEED TO IMPROVE MY BASIC SKILLS.	①
I AM TOO OLD TO GO BACK TO SCHOOL.	②
I THINK SCHOOL IS TOO HARD.	③
I DON'T HAVE THE TIME.	④
I DON'T LIKE SCHOOL.	⑤
I HAVE TOO MANY CONFLICTS.	⑥
IT WOULD TAKE TOO LONG TO FINISH A BASIC SKILLS PROGRAM.	⑦
I DON'T HAVE ANY INFORMATION AVAILABLE ABOUT BASIC SKILLS PROGRAMS.	⑧

IA-3. How likely is it that you will move out of Iowa in the next five years? Would you say…

not likely,	①
somewhat likely, or	②
very likely?	③

IA-4. Do you feel that a state's literacy rate affects an out-of-state employer's decision to establish a new location there?

YES	①
NO	②
NO OPINION	③

IA-5. Do you feel that an employer has an obligation to provide literacy education to its employees who need assistance?

YES	①
NO	②
NO OPINION	③

LOUISIANA

LA-1. For which of the following reasons would you enroll in a course or a training program? (CODE ALL THAT APPLY.)

To qualify for a better job ..... ①
To participate in leisure activities or for personal improvement ..... ②
To complete a diploma or degree, or to qualify for a license ..... ③
To improve job-related skills ..... ④
To improve parenting/child-care skills ..... ⑤
NONE OF THE ABOVE ..... ⑥

LA-2. HAND CARD: Which one of the following reasons would most likely keep you from enrolling in a course or a training program?

HAND CARD 13

LACK OF TIME OR INTEREST. ..... ①
COURSE OR TRAINING PROGRAM NOT OFFERED AT CONVENIENT TIMES. ..... ②
INABILITY TO PAY FOR COURSE OR TRAINING. ..... ③
LACK OF CHILD CARE. ..... ④
LACK OF TRANSPORTATION, OR ..... ⑤
LACK OF INFORMATION ABOUT AVAILABLE COURSES OR TRAINING PROGRAMS? ..... ⑥
NONE OF THE ABOVE ..... ⑦

LA-3. If you are currently employed, in which of the following areas do you feel you need additional training to become a more productive worker in your present job? (CODE ALL THAT APPLY.)

Reading ..... ①
Arithmetic ..... ②
Computers ..... ③
Problem solving ..... ④
Listening ..... ⑤
Working in teams ..... ⑥
None of the above ..... ⑦
NOT CURRENTLY EMPLOYED ..... ⑧

LA-4. Have you ever been passed over for a job or a promotion for any of the following reasons? (CODE ALL THAT APPLY.)

Not having a high school diploma ..... ①
Not speaking English well enough ..... ②
Not reading well enough ..... ③
Not writing well enough ..... ④
Not being able to do arithmetic well enough ..... ⑤
NEVER EMPLOYED ..... ⑥
NONE OF THE ABOVE ..... ⑦

LA-5. If you do not have a high school diploma or GED (general equivalency diploma), would you be willing to work toward obtaining a GED if your employer allowed you to attend classes during work hours?

YES ..... ①
NO ..... ②
NOT APPLICABLE ..... ③

NEW JERSEY

NJ-1. If you are currently employed, in which of the following areas do you feel you need additional training to become a more productive worker in your present job? (CODE ALL THAT APPLY.)

Reading ①
Arithmetic ②
Computers ③
Problem solving ④
Listening ⑤
Working in teams ⑥
None of the above ⑦
NOT CURRENTLY EMPLOYED ⑧

NJ-2. Have you ever been passed over for a job or a promotion for any of the following reasons? (CODE ALL THAT APPLY.)

Not having a high school diploma ①
Not speaking English well enough ②
Not reading well enough ③
Not writing well enough ④
Not being able to do arithmetic well enough ⑤
NEVER EMPLOYED ⑥
NONE OF THE ABOVE ⑦

NJ-3. HAND CARD: Which one of the following reasons would most likely keep you from enrolling in a course or a training program?

HAND CARD 13

LACK OF TIME OR INTEREST, ①
COURSE OR TRAINING PROGRAM NOT OFFERED AT CONVENIENT TIMES, ②
INABILITY TO PAY FOR COURSE OR TRAINING, ③
LACK OF CHILD CARE, ④
LACK OF TRANSPORTATION, OR ⑤
LACK OF INFORMATION ABOUT AVAILABLE COURSES OR TRAINING PROGRAMS? ⑥
NONE OF THE ABOVE ⑦

NJ-4. HAND CARD: Where would the best location be for you to take a course or a training program?

HAND CARD 12

A LOCAL PUBLIC SCHOOL OUTSIDE OF SCHOOL HOURS, ①
A LOCAL COMMUNITY COLLEGE, ②
A UNIVERSITY, ③
A LIBRARY, CHURCH, COMMUNITY CENTER, OR ADULT EDUCATION CENTER, ④
YOUR WORKPLACE, OR ⑤
SOME OTHER LOCATION? ⑥
NOT INTERESTED IN TAKING A COURSE ⑦

NJ-5. What type of adult education program would best suit your needs?

A program leading to a credential, certificate, or license ①
A program that teaches parenting skills ②
A program that teaches work skills at the work site ③
A program for self-enrichment ④
A program that combines both basic skills and job training ⑤
NONE OF THE ABOVE ⑥

NEW YORK

NY-1. How many years have you lived in New York?

LESS THAN ONE YEAR .... ①
1 TO 5 YEARS .... ②
6 TO 10 YEARS .... ③
11 TO 15 YEARS .... ④
16 TO 20 YEARS .... ⑤
MORE THAN 20 YEARS .... ⑥

NY-2. Which of the following statements apply to you? (CODE ALL THAT APPLY.)

I grew up in a home where there were a variety of books, newspapers, magazines, and other printed materials. .... ①
I was read to as a young child by my parent(s) or guardian(s). .... ②
I was helped with schoolwork by my parent(s) or guardian(s). .... ③
NONE OF THE ABOVE .... ④

NY-3. What type of adult education program would best suit your needs?

A program leading to a credential, certificate, or license .... ①
A program that teaches parenting skills .... ②
A program that teaches work skills at the work site .... ③
A program for self-enrichment .... ④
A program that combines both basic skills and job training .... ⑤
NONE OF THE ABOVE .... ⑥

NY-4. Has lack of adequate child care ever prevented you from taking an adult literacy or job training program?

YES .... ①
NO .... ②
NO OPINION .... ③

NY-5. Has lack of convenient transportation ever prevented you from taking an adult literacy or job training program?

YES .... ①
NO .... ②
NO OPINION .... ③

OHIO

OH-1. How many years have you lived in Ohio?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

OH-2. Which of the following levels of schooling did you attend in Ohio? (CODE ALL THAT APPLY.)

Kindergarten through 3rd grade	①
Grades 4 through 8	②
Grades 9 through 12	③
Vocational school or community college (One- or two-year program)	④
Four-year college or university	⑤
Did not attend school in Ohio	⑥

OH-3. If you moved into Ohio within the last five years, what was your primary reason for moving?

DID NOT MOVE INTO OHIO WITHIN THE LAST 5 YEARS	①
ACCOMPANIED FAMILY OR ANOTHER PERSON	②
TO FIND A JOB	③
TO ACCEPT A NEW JOB OR TRANSFER	④
TO GO TO SCHOOL	⑤
SOME OTHER REASON	⑥

OH-4. If you received a high school diploma, where did you receive it? Did you receive it from…

a public school in Ohio,	①	(STOP HERE; END OF OHIO QUESTIONS)
a public school outside Ohio,	②	
a private school in Ohio, or	③	
a private school outside Ohio?	④	
DID NOT RECEIVE A HIGH SCHOOL DIPLOMA	⑤	(OH-5)

OH-5. If you received a certificate of high school equivalence as a result of taking and passing the General Educational Development (GED) tests, where did you receive it? Did you receive a GED…

In Ohio, or	①
Outside Ohio?	②
DID NOT RECEIVE A GED	③

PENNSYLVANIA

PA-1. How many years have you lived in Pennsylvania?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

PA-2. How likely is it that you will move out of Pennsylvania in the next five years? Would you say...

not likely,	①
somewhat likely, or	②
very likely?	③

PA-3. Which of the following statements apply to you? (CODE ALL THAT APPLY.)

I grew up in a home where there were a variety of books, newspapers, magazines, and other printed materials.	①
I was read to as a young child by my parent(s) or guardian(s).	②
I was helped with schoolwork by my parent(s) or guardian(s).	③
NONE OF THE ABOVE	④

PA-4. Do you feel that an employer has an obligation to provide literacy education to its employees who need assistance?

YES	①
NO	②
NO OPINION	③

PA-5. If you do not have a high school diploma or GED (general equivalency diploma), would you be willing to work toward obtaining a GED if your employer allowed you to attend classes during work hours?

YES	①
NO	②
NOT APPLICABLE	③

TEXAS

TX-1. How many years have you lived in Texas?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

TX-2. Which of the following levels of schooling did you attend in Texas? (CODE ALL THAT APPLY.)

Kindergarten through 3rd grade	①
Grades 4 through 8	②
Grades 9 through 12	③
Vocational school or community college (One- or two-year program)	④
Four-year college or university	⑤
Did not attend school in Texas	⑥

TX-3. Which of the following statements apply to you? (CODE ALL THAT APPLY.)

I grew up in a home where there were a variety of books, newspapers, magazines, and other printed materials.	①
I was read to as a young child by my parent(s) or guardian(s).	②
I was helped with schoolwork by my parent(s) or guardian(s).	③
NONE OF THE ABOVE	④

TX-4. If you have any children under six years of age living with you now, how often do you read to or with them?

EVERY DAY	①
A FEW TIMES A WEEK	②
ABOUT ONCE A WEEK	③
ABOUT ONCE OR TWICE A MONTH	④
ALMOST NEVER	⑤
RESPONDENT HAS NO CHILDREN UNDER SIX YEARS OF AGE IN THE HOUSEHOLD.	⑥

TX-5. HANDCARD: Which one of the following reasons would most likely keep you from enrolling in a course or a training program?

HAND CARD 13

LACK OF TIME OR INTEREST,	①
COURSE OR TRAINING PROGRAM NOT OFFERED AT CONVENIENT TIMES,	②
INABILITY TO PAY FOR COURSE OR TRAINING,	③
LACK OF CHILD CARE,	④
LACK OF TRANSPORTATION, OR	⑤
LACK OF INFORMATION ABOUT AVAILABLE COURSES OR TRAINING PROGRAMS?	⑥
NONE OF THE ABOVE	⑦

WASHINGTON

WA-1. How many years have you lived in Washington?

LESS THAN ONE YEAR	①
1 TO 5 YEARS	②
6 TO 10 YEARS	③
11 TO 15 YEARS	④
16 TO 20 YEARS	⑤
MORE THAN 20 YEARS	⑥

WA-2. Which of the following levels of schooling did you attend in Washington? (CODE ALL THAT APPLY.)

Kindergarten through 3rd grade	①
Grades 4 through 8	②
Grades 9 through 12	③
Vocational school or community college (One- or two-year program)	④
Four-year college or university	⑤
Did not attend school in Washington	⑥

WA-3. HANDCARD: Which one of the following reasons would most likely keep you from enrolling in a course or a training program?

HAND CARD 13

LACK OF TIME OR INTEREST,	①
COURSE OR TRAINING PROGRAM NOT OFFERED AT CONVENIENT TIMES,	②
INABILITY TO PAY FOR COURSE OR TRAINING,	③
LACK OF CHILD CARE,	④
LACK OF TRANSPORTATION, OR	⑤
LACK OF INFORMATION ABOUT AVAILABLE COURSES OR TRAINING PROGRAMS?	⑥
NONE OF THE ABOVE	⑦

WA-4. If you are currently employed, in which of the following areas do you feel you need additional training to become a more productive worker in your present job? (CODE ALL THAT APPLY.)

Reading	①
Arithmetic	②
Computers	③
Problem solving	④
Listening	⑤
Working in teams	⑥
None of the above	⑦
NOT CURRENTLY EMPLOYED	⑧

WA-5. Have you ever been passed over for a job or a promotion for any of the following reasons? (CODE ALL THAT APPLY.)

Not having a high school diploma	①
Not speaking English well enough	②
Not reading well enough	③
Not writing well enough	④
Not being able to do arithmetic well enough	⑤
NEVER EMPLOYED	⑥
NONE OF THE ABOVE	⑦

Appendix H

ENGLISH BACKGROUND QUESTIONNAIRE FOR PRISONS

USING PRINTED MATERIALS

A SURVEY OF AMERICA'S ADULTS

NALS

CAREFULLY PLACE LABEL INSIDE THIS BOX

INCARCERATED ENGLISH BACKGROUND QUESTIONNAIRE

Booklet ID #	Office Use Only	BQ Elapsed Time (minutes)	EX Elapsed Time (minutes)

The National Center for Education Statistics is supporting the conduct of NALS through a contract awarded to

Educational Testing Service

Copyright © 1991 by Educational Testing Service. All Rights Reserved.

OMB No. 1850-0654. Approval Expires August 1993

Public reporting burden for this collection of information is estimated to average 1.17 hours per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information; including suggestions for reducing this burden, to the U.S. Department of Education, Information Management and Compliance Division, Washington, D.C. 20202-4651; and to the Office of Management and Budget, Paperwork Reduction Project 1850-NEW, Washington, D.C. 20503.

A Survey of America's Adults Incarcerated Version

BACKGROUND QUESTIONNAIRE

Hello I am (NAME) from Westat, a research organization in Rockville, MD. We are helping the United States Department of Education with a very important survey about how adults use printed materials. Based on a list of all persons residing in this institution, you have been selected at random to take part in a special version of this survey designed to collect information from persons incarcerated in state and federal prisons.

I will ask you a short set of questions about your educational experiences, work history, and reading and writing activities. Then, I will give you a booklet containing exercises based on printed materials such as newspapers, maps, stories, forms, and advertisements. Others who have completed these exercises found them interesting and fun. The entire survey will take approximately one hour to complete [and your personal "canteen" account will be credited $20 upon your completion of the interview and exercises.]

Your participation in this survey is voluntary and very important. All of your answers will be kept strictly confidential and will not be available to anyone here in this facility. Your responses will be combined with the responses of others to make totals and averages and will never be linked to your name.

IF RESPONDENT REQUESTS MORE INFORMATION ABOUT PURPOSE OF SURVEY: Today, adults of all ages are expected to use printed information in new and more complex ways, but there is very little information available on whether or not they are well prepared to do so. This survey will provide information about the reading and writing experiences, activities, and skills of adults in the United States. Educators, policymakers, and business leaders will use information from this study to help develop programs that will better equip us with the skills we will need for the future.

INTERVIEWER NAME: ______________________________

Scorer #1	Scorer #2
☐☐	☐☐

START
TIME: _____ A.M. _____ P.M.

SECTION A
GENERAL AND LANGUAGE BACKGROUND

A-1. In what country were you born?

USA (50 STATES OR D.C.)	1	**(A-4)**	
U.S. TERRITORY (SPECIFY): ____________________	2	**(A-2)**	
OTHER (SPECIFY COUNTRY): ____________________	3	**(A-2)**	_____

A-2. How many years have you lived in the United States?

1 TO 5	1
6 TO 10	2
11 TO 15	3
16 TO 20	4
21 TO 30	5
31 TO 40	6
41 TO 50	7
51 OR MORE	8

A-3. What was the highest level of education you completed before coming to the United States? (IF RESPONSE DOES NOT FIT CATEGORIES, PROBE FOR EQUIVALENT.)

DID NOT ATTEND SCHOOL BEFORE COMING TO U.S.	1
PRIMARY (GRADES K-3)	2
ELEMENTARY (GRADES 4-8)	3
SECONDARY (GRADES 9-12)	4
VOCATIONAL (POSTSECONDARY)	5
COLLEGE/UNIVERSITY	6
OTHER (SPECIFY): ____________________	7

A-4. When you were growing up, what language or languages were usually spoken in your home? (CODE ALL THAT APPLY.)

ENGLISH	1	
SPANISH	2	
OTHER (SPECIFY): ____________________	3	_____

A-5. What language or languages did you learn to speak before you started school? (CODE ALL THAT APPLY.)

ENGLISH	1
SPANISH	2
OTHER (SPECIFY): ______	3

BOX 1. IF ENGLISH ONLY FOR BOTH A-4 AND A-5, SKIP TO A-16; OTHERWISE, CONTINUE.

A-6. What language did you first learn to read and write?

ENGLISH	1
SPANISH	2
OTHER (SPECIFY): ______	3
NEVER LEARNED TO READ AND WRITE	4

A-7. How old were you when you learned to speak English?

1-4 YEARS	1
5-10 YEARS	2
11-15 YEARS	3
16-20 YEARS	4
21 OR OLDER	5
DOES NOT SPEAK ENGLISH	6

BOX 2. IF ENGLISH ONLY IN A-5, SKIP TO A-16; OTHERWISE, CONTINUE. IF MORE THAN ONE NON-ENGLISH LANGUAGE IN A-5, USE FIRST ONE MENTIONED FOR A-8 AND A-9.

A-8. With regard to (NON-ENGLISH LANGUAGE IN A-5), how well do you…

Would you say…

	very well	well	not well	or	not at all
understand it when it is spoken to you?	1	2	3		4
speak it?	1	2	3		4
read it?	1	2	3		4
write it?	1	2	3		4

A-9. HAND CARD: With regard to (NON-ENGLISH LANGUAGE IN A-5), how often do you…

HAND CARD 1

Would you say…

	every day	a few times a week	once a week	less than once a week	or never
A. listen to radio programs, tapes, or records in (LANGUAGE)?	1	2	3	4	5
B. watch television programs or video tapes in (LANGUAGE)?	1	2	3	4	5
C. read newspapers, magazines, or books in (LANGUAGE)?	1	2	3	4	5
D. write or fill out letters or forms in (LANGUAGE)?	1	2	3	4	5

A-10. Have you ever taken a course to learn how to read and write English as a second language?

YES	1	**(A-11)**
NO	2	**(A-12)**

A-11. Did you complete this course?

YES	1
NO	2

A-12. Have you ever taken a course to learn how to speak and understand English as a second language?

YES	1	**(A-13)**
NO	2	**(A-14)**

A-13. Did you complete this course?

YES	1
NO	2

A-14. Which language do you usually speak now?

ENGLISH	1
SPANISH	2
OTHER (SPECIFY): ________________	3

A-15. What other language do you often speak now?

ENGLISH	1	
SPANISH	2	
OTHER (SPECIFY): ______________________	3	____
NO OTHER LANGUAGE SPOKEN	4	

A-16. With regard to the English language, how well do you...

Would you say...

	very well	well	not well	or	not at all
understand it when it is spoken to you?	1	2	3		4
speak it?	1	2	3		4
read it?	1	2	3		4
write it?	1	2	3		4
Do arithmetic problems when you have to get the numbers from written materials?	1	2	3		4

SECTION B
EDUCATIONAL BACKGROUND AND EXPERIENCES

B-1. HAND CARD: I'd like to ask you about your educational background and experiences. What is the highest level of public or private education you completed in the United States? (IF R WENT TO SCHOOL OUTSIDE U.S., PROBE FOR EQUIVALENT.)

HAND CARD 2

A.	*(NOT APPLICABLE TO INCARCERATED SURVEY)*		
B.	LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE): _____	02	(B-2)
C.	SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT COMPLETE 12TH GRADE) (SPECIFY GRADE):_____	03	(B-2)
D.	GED OR HIGH SCHOOL EQUIVALENCY	04	(B-2)
E.	HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM)	05	(B-4)
F.	ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL	06	
G.	COLLEGE: LESS THAN TWO YEARS	07	
H.	COLLEGE: ASSOCIATE'S DEGREE (A.A.)	08	
I.	COLLEGE: TWO YEARS OR MORE, NO DEGREE	09	
J.	COLLEGE GRADUATE (B.S. OR B.A.)	10	
K.	POSTGRADUATE/NO DEGREE	11	
L.	POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.)	12	
M.	DON'T KNOW	13	(B-5)

B-2. What was the main reason you stopped your public or private schooling when you did?

FINANCIAL PROBLEMS	1
WENT TO WORK OR INTO THE MILITARY	2
PREGNANCY	3
LOST INTEREST OR BEHAVIOR PROBLEMS IN SCHOOL	4
ACADEMIC PROBLEMS AT SCHOOL	5
FAMILY OR PERSONAL PROBLEMS	6
CONVICTED OF CRIME OR SENT TO JAIL/ PRISON/DETENTION CENTER	7
OTHER	8

BOX 3. CHECK B-1. IF CODED 04, SKIP TO B-5; OTHERWISE, CONTINUE.

B-3. Did you ever study for a GED or high school equivalency?

YES	1
NO	2

BOX 4. CHECK B-1. IF CODED 02, SKIP TO B-5; OTHERWISE, CONTINUE.

B-4. What kind of high school program did you take? Was it...

college preparatory?	1
vocational, technical, or trade?	2
general?	3
or the same for everyone?	4

B-5. Since your current admission to prison, have you ever been in any education program, excluding vocational training?

YES	1	(B-6)
NO	2	(B-9)

B-6. What kind of program was that — basic classes up to the 9th grade, high school classes to get a diploma or GED, or college-level classes? (CODE ALL THAT APPLY.)

BASIC CLASSES UP TO 9TH GRADE	1
HIGH SCHOOL CLASSES FOR DIPLOMA OR GED	2
COLLEGE-LEVEL CLASSES	3
OTHER	4

B-7. Are you still in an education program?

YES	1
NO	2

B-8. Since your current admission to prison, how long have you spent altogether in education program(s)?

LESS THAN ONE MONTH	1
NUMBER OF MONTHS (SPECIFY): ____________	2
NUMBER OF YEARS (SPECIFY): ____________	3
DON'T KNOW	4

B-9. Since your current admission to prison, have you ever been in any vocational training program, excluding prison work assignments?

YES .... 1 **(B-10)**
NO .... 2 **(B-12)**

B-10. Are you still in a vocational program?

YES 1
NO 2

B-11. Since your current admission to prison, how long have you spent altogether in vocational training?

LESS THAN ONE MONTH 1
NUMBER OF MONTHS (SPECIFY): ____________________ 2
NUMBER OF YEARS (SPECIFY): ____________________ 3
DON'T KNOW 4

B-12. Before your current admission to prison, did you ever take part in a program other than in regular school in order to improve your *basic skills*, that is, basic reading, writing, and arithmetic skills?

YES .... 1 **(B-13)**
NO .... 2 **(B-15)**

ASK B-13 AND B-14 IN SEQUENCE FOR EACH PROGRAM CODED "YES" IN B-13.

B-13. Was this program…	YES	NO	DON'T KNOW	**B-14.** IF YES: How recently did you take part in this training?
a training program or courses given or sponsored by your employer or union?	1	2	3	WITHIN THE PAST YEAR ...... 2 BETWEEN ONE & FIVE YEARS ...... 3 MORE THAN FIVE YEARS AGO 4
a publicly sponsored education and training program, such as JTPA or ABE?	1	2	3	WITHIN THE PAST YEAR ...... 2 BETWEEN ONE & FIVE YEARS ...... 3 MORE THAN FIVE YEARS AGO 4
a tutoring program sponsored by a library, church, or community organization?	1	2	3	WITHIN THE PAST YEAR ...... 2 BETWEEN ONE & FIVE YEARS ...... 3 MORE THAN FIVE YEARS AGO 4
any other program (such as one offered by the military, prisons, or other institutions)?	1	2	3	WITHIN THE PAST YEAR ...... 2 BETWEEN ONE & FIVE YEARS ...... 3 MORE THAN FIVE YEARS AGO 4

B-15. Do you have a physical, mental, or other health condition that keeps you from participating fully in work, school, or other activities?

YES	1
NO	2

B-16. Do you have any difficulty seeing the words and letters in ordinary newspaper print even when wearing glasses or contact lenses, if you usually wear them?

YES	1
NO	2

B-17. Do you have any difficulty hearing what is said in a normal conversation with another person even when using a hearing aid, if you usually wear one?

YES	1
NO	2

B-18. Do your currently have any of these conditions? (READ LIST AND CODE YES OR NO FOR EACH CONDITION.)

	YES	NO
A learning disability	1	2
Any mental or emotional condition	1	2
Mental retardation	1	2
A speech disability	1	2
A physical disability	1	2
A long-term illness (6 months or more)	1	2
Any other health impairment	1	2

SECTION C
CURRENT OFFENSES AND CRIMINAL HISTORY

C-1. For what offenses are you now in prison? (PROBE: Any others?) (RECORD EACH OFFENSE ON A SEPARATE LINE BELOW.)

OFFENSE #1: ____________________
OFFENSE #2: ____________________
OFFENSE #3: ____________________
OFFENSE #4: ____________________
OFFENSE #5: ____________________ ____

BOX 5. CHECK C-1. IF ONLY OFFENSE RECORDED, SKIP TO C-3; OTHERWISE, CONTINUE.

C-2. For which of these offenses did you receive the longest sentence?

____________________ ____

C-3. In what month and year were you first arrested for (any of these offenses/this offense)?

____________________ ____
MONTH / YEAR

DON'T KNOW	1	

C-4. In what month and year were you first admitted to prison for (these offenses/this offense)?

____________________ ____
MONTH / YEAR

DON'T KNOW	1	

C-5. Have you ever been placed on probation, either as a juvenile or as an adult?

YES	1	(C-6)
NO	2	(C-8)

C-6. How many times have you ever been placed on probation as a juvenile?

NONE	1	
NUMBER OF TIMES (SPECIFY): ____________	2	____
DON'T KNOW	3	

C-7. How many times have you been placed on probation as an adult?

NONE 1
NUMBER OF TIMES (SPECIFY): ________ .... 2 ____
DON'T KNOW 3

C-8. Did you ever serve time in prison, jail, or some other correctional facility for any other offenses?

YES 1 (C-9)
NO 2 (C-11)

C-9. How many times as a juvenile?

NONE 1
NUMBER OF TIMES (SPECIFY): ________ .... 2 ____
DON'T KNOW 3

C-10. How many times as an adult?

NONE 1
NUMBER OF TIMES (SPECIFY): ________ .... 2 ____
DON'T KNOW 3

C-11. Do you have a define date on which you expect to be released from prison?

YES 1 (C-12)
NO 2 (C-13)

C-12. In what month and year will you be released?

________ → (D-1)
MONTH / YEAR ____

C-13. Do you expect to ever be released from prison?

YES 1 (C-14)
NO 2 (D-1)

C-14. In what month and year is your earliest possible release date?

________ ____
MONTH / YEAR

DON'T KNOW 1

SECTION D
PRISON WORK ASSIGNMENTS AND LABOR FORCE PARTICIPATION

D-1. Now I'd like to ask you some questions about work assignments. Do you currently have any work assignments either inside or outside of this prison facility?

YES	1	(D-2)
NO	2	(D-8)

D-2. What work assignments do you have now? (CODE ALL THAT APPLY.)

GOODS PRODUCTION (TAG SHOP, PRINT SHOP, ETC.)	01
GENERAL JANITORIAL DUTIES (CLEANING)	02
GROUNDS OR ROAD MAINTENANCE	03
FOOD PREPARATION OR RELATED DUTIES (KITCHEN, BAKERY, BUTCHERY, ETC.)	04
LAUNDRY	05
HOSPITAL, INFIRMARY, OR OTHER MEDICAL SERVICES	06
FARMING/FORESTRY/RANCHING	07
OTHER SERVICES (LIBRARY, STOCROOM, STORE, OFFICE HELP, RECREATION, BARBER OR BEAUTY SHOP, ETC.)	08
MAINTENANCE OR REPAIR/CONSTRUCTION	09
ENROLLED IN SCHOOL	10
OTHER (SPECIFY): ____________________	11

D-3. In the last week, how many hours were you assigned to (this job/these jobs)?

HOURS (SPECIFY): __________ ____

D-4. Are you paid money for any of this work?

YES	1	(D-5)
NO	2	(D-6)

D-5. How much did you make last month (from all work assignments)? (WRITE IN DOLLAR AMOUNT AND CODE APPROPRIATE UNIT.)

$________ . ________ ____

PER HOUR	1
PER DAY	2
PER WEEK	3
PER TWO-WEEK PERIOD	4
PER MONTH	5
PER YEAR	6
OTHER (SPECIFY): ____________________	7

D-6. (Other than money,) do you receive anything for work, such as time credits or other privileges?

YES	1	(D-7)
NO	2	(D-8)

D-7. What do you receive? (CODE ALL THAT APPLY.)

GOOD TIME	1
CIGARETTES, FOOD, ETC.	2
EXTRA PRIVILEGES	3
OTHER (SPECIFY): ______	4

D-8. HAND CARD: In the year before your arrest on (READ ARREST DATE FROM QUESTION C-3 ON PAGE 12), did you receive income from any of the following? (CODE ALL THAT APPLY.)

HAND CARD 3

A. Social Security or other pensions (e.g., veterans, Supplemental Security Income, disability)	1
B. Welfare, charity	2
C. Unemployment insurance compensation and/or workman's compensation	3
D. Family or friends	4
E. Illegal sources	5
F. Educational scholarship or grant	6
G. Anything else? (SPECIFY): ______	7
H. NONE	8

BOX 6. CHECK ADMISSION DATE FROM QUESTION C-4 ON PAGE 12. IF YEAR OF ADMISSION IS 1988 OR EARLIER, SKIP TO E-1;OTHERWISE, CONTINUE.

D-9. In the past three years, that is, since January 1989, did you work for pay or profit, either full-time or part-time, while not serving time in prison?

YES	1	(D-10)
NO	2	(E-1)

D-10. Between January 1989 and (READ ADMISSION DATE FROM QUESTION C-4 ON PAGE 12), how many months did you work?

LESS THAN 1 MONTH	1	(E-1)
NUMBER OF MONTHS (SPECIFY): ______	2	(D-11) ______

D-11. Now I would like to ask you some questions about your most recent full-time or part-time job prior to your current admission to prison. For what kind of business or industry did you work? (For example, television and radio manufacturing, retail shoe store, state labor department, farm, etc.) (IF R WORKED TWO OR MORE JOBS, PROBE: Tell me about the job you worked the most hours or the job you considered your primary employment.)

BUSINESS OR INDUSTRY: ______________________________

____ ____

D-12. What was your occupation, that is, what was your job called? (For example, electrical engineer, stock clerk, typist, farmer, etc.)

WORK: ____

____ ____

D-13. What were the most important activities or duties at this job? (For example, typing, keeping account books, filing, selling cars, operating a printing press, finishing concrete, etc.)

ACTIVITIES OR DUTIES: ______________________________

____

____

____

D-14. Was this full-time, part-time, or occasional work?

FULL-TIME 1
PART-TIME 2
OCCASIONAL 3

D-15. HAND CARD: Which category on this card represents your average monthly earnings for that job before any deductions?

HAND CARD 4

A. $1- $99	01
B. $100- $199	02
C. $200- $299	03
D. $300- $399	04
E. $400- $499	05
F. $500- $599	06
G. $600- $799	07
H. $800- $999	08
I. $1,000- $1,199	09
J. $1,200- $1,499	10
K. $1,500- $1,999	11
L. $2,000- $2,499	12
M. $2,500- $4,999	13
N. $5,000 OR MORE	14
O. DON'T KNOW	15
P. NO INCOME	16

SECTION E
LITERACY ACTIVITIES AND COLLABORATION

E-1. Now I'd like to talk to you about what you read in English. First, let's talk about newspapers. How often do you read a newspaper — every day, a few times a week, once a week, less than once a week, or never?

EVERY DAY	1	(E-2)
A FEW TIMES A WEEK	2	(E-2)
ONCE A WEEK	3	(E-2)
LESS THAN ONCE A WEEK	4	(E-2)
NEVER	5	(E-3)

E-2. Now I will read a list of the different parts of newspapers. Would you please tell me which parts you generally read when looking at a newspaper? Do you read… (CODE ALL THAT APPLY.)

news?	01
sports?	02
home, fashion, or health?	03
editorial pages?	04
financial news or stock listings?	05
comics?	06
classified ads or other advertisements?	07
TV, movie, or concert listings?	08
book, movie, or art reviews?	09
horoscope or advice columns?	10
something else? (SPECIFY): ______________________	11
NONE?	12

E-3. Next, let's talk about magazines. About how many different magazines do you look at or read in English on a regular basis?

0	1
1	2
2	3
3 TO 5	4
6 OR MORE	5

E-4. Next, I'd like to ask you about books you may have read in English recently. I will read a list of different types of books. Would you please tell me which of these types of books you've read in English in the past six months, if any? Have you read… (CODE ALL THAT APPLY.) (INCLUDE COURSE BOOKS.)

fiction?	1
recreation or entertainment?	2
current affairs or history?	3
inspiration or religion?	4
science or social science?	5
reference, such as encyclopedias or dictionaries?	6
manuals for cooking, operating, repairing, or building?	7
any other types of books? (SPECIFY): ____________	8
NONE?	9

E-5. I would like to ask you about other kinds of things you may read in English. Please consider printed or written materials, as well as information displayed on computer screens. HAND CARD: How often do you read or use information from (ITEM)? Would you say every day, a few times a week, once a week, less than once a week, or never?

HAND CARD 1

	EVERY DAY	A FEW TIMES A WEEK	ONCE A WEEK	LESS THAN ONCE A WEEK	NEVER
A. Letters or memos	1	2	3	4	5
B. Reports, articles, magazines, or journals	1	2	3	4	5
C. Manuals or reference books, including catalogs or parts lists	1	2	3	4	5
D. Directions or instructions for medicines, recipes, or other products	1	2	3	4	5
E. Diagrams or schematics	1	2	3	4	5
F. Bills, invoices, spreadsheets, or budget tables	1	2	3	4	5

E-6. Now I would like to find out about things that you write or fill out in English. Please consider writing or typing on paper, as well as entering information into computers. HAND CARD: How often do you write or fill out (ITEM)? Would you say every day, a few times a week, once a week, less than once a week, or never?

HAND CARD 1

	EVERY DAY	A FEW TIMES A WEEK	ONCE A WEEK	LESS THAN ONCE A WEEK	NEVER
A. Letters or memos	1	2	3	4	5
B. Forms	1	2	3	4	5
C. Reports or articles	1	2	3	4	5

E-7. How often do you use arithmetic or mathematics, that is, add, subtract, multiply, divide, or measure? Would you say every day, a few times a week, once a week, less than once a week, or never?

EVERY DAY	1
A FEW TIMES A WEEK	2
ONCE A WEEK	3
LESS THAN ONCE A WEEK	4
NEVER	5

E-8. HAND CARD: How much help do you get from family members, friends, or other people with…

HAND CARD 5

Would you say…

	a lot	some	a little	or	none
A. filling out forms?	1	2	3		4
B. reading or explaining newspaper articles or other written information?	1	2	3		4
C. printed information associated with government agencies, public companies, private businesses, hospitals, etc.?	1	2	3		4
D. writing notes and letters?	1	2	3		4
E. using basic arithmetic, that is, adding, subtracting, multiplying, or dividing, such as in filling out order forms?	1	2	3		4

E-9. Since your admission to prison on (READ ADMISSION DATE FROM QUESTION C-4 ON PAGE 12), have you joined any organization authorized by prison authorities, such as a prison religious group, a study group, a prisoner assistance group, or other self-help group?

YES	1	(E-10)
NO	2	(F-1)

E-10. What groups have you joined? (CODE ALL THAT APPLY.)

DRUG AWARENESS OR OTHER DRUG DEPENDENCY GROUP	01
ALCOHOLICS ANONYMOUS, AL-ANON, OR OTHER ALCOHOL-RELATED GROUP	02
A BIBLE CLUB OR OTHER RELIGIOUS STUDY GROUP (INCLUDING MUSLIMS)	03
AN ETHNIC/RACIAL ORGANIZATION (E.G., NAACP, AFRICAN AMERICAN/BLACK CULTURE GROUP, HISPANIC COMMITTEE, AZTLAN, OR LAKOTA)	04
PRISONER ASSISTANCE GROUPS (E.G., INMATE LIAISON, ADVISORY, OR WORKER'S COUNCILS) OR PRISONER COUNSELING GROUP	05
OTHER PRISONER SELF-HELP/PERSONAL IMPROVEMENT GROUP (E.G., LIFER GROUP, PARENT AWARENESS GROUP, TOASTMASTERS, JAYCEES, GAVEL CLUB, VETERANS CLUB)	06
OUTSIDE COMMUNITY ACTIVITIES	07
RELIGIOUS ACTIVITIES	08
CLASSES IN PARENTING/CHILDREARING SKILLS	09
CLASSES IN LIFE SKILLS (INCLUDING HOUSEHOLD FINANCE, HOW TO FIND A JOB, ETC.)	10
CLASSES IN ARTS, CRAFTS	11
PRERELEASE PROGRAMS	12
OTHER (SPECIFY): ______________________________	13

SECTION F
DEMOGRAPHIC INFORMATION

F-1. When you were growing up, who did you live with most of the time? (IF MORE THAN ONE RELATIVE, CODE CLOSET BLOOD RELATIVE.)

MOTHER	1	(F-2)
FATHER	2	(F-3)
BOTH PARENTS (INCLUDING ONE STEPPARENT)	3	(F-2)
GRANDPARENT(S)	4	(F-2)
OTHER RELATIVE(S) (INCLUDING STEPRELATION(S))	5	(F-2)
FRIEND(S)	6	(F-4)
FOSTER HOME(S)	7	(F-4)
AGENCY OR INSTITUTION (INCLUDING RELIGIOUS INSTITUTIONS)	8	(F-4)
SOMEONE ELSE (SPECIFY): ______________________	9	(F-4)

F-2. What was the highest level of education your mother (stepmother or female guardian) completed? (IF WENT TO SCHOOL OUTSIDE U.S., PROBE FOR EQUIVALENT.)

LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE): ________	01
SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT COMPLETE 12TH GRADE) (SPECIFY GRADE): ________	02
GED OR HIGH SCHOOL EQUIVALENCY	03
HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM)	04
ATTENDED A VOCATIONAL, TRADE OR BUSINESS SCHOOL AFTER HIGH SCHOOL	05
COLLEGE: LESS THAN TWO YEARS	06
COLLEGE: ASSOCIATE'S DEGREE (A.A.)	07
COLLEGE: TWO YEARS OR MORE, NO DEGREE	08
COLLEGE GRADUATE (B.S. OR B.A.)	09
POSTGRADUATE/NO DEGREE	10
POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.)	11
DON'T KNOW	12

______

BOX 7. CHECK F-1. IF CODED 1, GO TO F-4; OTHERWISE, CONTINUE.

F-3. What was the highest level of education your father (stepfather or male guardian) completed? (IF WENT TO SCHOOL OUTSIDE U.S., PROBE FOR EQUIVALENT.)

LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE): ________	01
SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT COMPLETE 12TH GRADE) (SPECIFY GRADE): ________	02
GED OR HIGH SCHOOL EQUIVALENCY	03
HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM)	04
ATTENDED A VOCATIONAL, TRADE OR BUSINESS SCHOOL AFTER HIGH SCHOOL	05
COLLEGE: LESS THAN TWO YEARS	06
COLLEGE: ASSOCIATE'S DEGREE (A.A.)	07
COLLEGE: TWO YEARS OR MORE, NO DEGREE	08
COLLEGE GRADUATE (B.S. OR B.A.)	09
POSTGRADUATE/NO DEGREE	10
POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.)	11
DON'T KNOW	12

______

F-4. What is your current marital status?

Never married	1
Married	2
Separated or divorced	3
Widowed	4

F-5. HAND CARD: Which of the groups on this card best describes you? (IF RESPONDENT REFUSES TO ANSWER, PLEASE RECORD THE RESPONDENT'S RACE FROM OBSERVATION.)

HAND CARD 6

A. WHITE	1
B. BLACK (AFRICAN AMERICAN)	2
C. AMERICAN INDIAN	3
D. ALASKAN NATIVE	4
E. PACIFIC ISLANDER	5
F. ASIAN (SPECIFY): ________	6
G. OTHER (SPECIFY): ________	7

F-6. Are you of Spanish or Hispanic origin or descent?

YES	1	(F-7)
NO	2	(F-8)

F-7. *HAND CARD: Which of the groups on this card best describes your Hispanic origin?*

HAND CARD 7

A. MEXICANO, MEXICAN, MEXICAN AMERICAN, CHICANO	1
B. PUERTO RICAN	2
C. CUBAN	3
D. CENTRAL/SOUTH AMERICAN	4
E. OTHER SPANISH/HISPANIC (SPECIFY): ______________________	5

F-8. What is your date of birth?

____ ____

MONTH / DAY / YEAR

INTERVIEWER PLEASE NOTE

F-9. SEX

MALE	1
FEMALE	2

END
TIME: ____ A.M. _____ P.M.

Appendix I

DERIVED VARIABLES

Appendix I: Derived Variables
NATIONAL ADULT LITERACY SURVEY

Listed below are all the derived variables for the 1992 National Adult Literacy Survey for version 2.

FOR HOUSEHOLD, PRISONER, AND NON-INCENTIVE SAMPLES	FOR PRISONER SAMPLE
Dage	Dib0801
Hhsize *	Dib1101
Drace	Dic0301
Dcbirth	Dic0401
Dagearr	Dic0601
Dlanghm	Dic0701
Dlangbs	Dic0901
Dlangrw	Dic1001
Dlangsp	Dic1201
Dbiling	Dic1401
Dbilit	Dd03601
Dlingl	Did0501
Dliter	Did1001
Dweekwg *	
Dwkyrwg *	
Dfaminc *	
Dbusind	
Doccup	
Dwkswrk *	
Clfstat *	
Annearn *	
Labfres *	
Hedenrn *	
Bascany *	
Bascone *	
Basefiv *	
Yrslfex *	
Povlevl *	
Sampcod	
Idcode	
Niruse	

* Derived variables not included in the prison dataset.

DAGE (ACTUAL AGE IN YEARS: BIRTH DATE F-12 MM/DD/YY)

Test date of 9/1/92 minus Birth date

TDATE = 365 x 92 + 30.4374 x 9 + 1
BDATE = 365 x YY + 30.4374 x MM + DD

AGEDAYS = TDATE - BDATE
AGEMONTH = AGEDAYS / 30.4374
IYEARS = AGEMONTH / 12

Adjustments for respondents born prior to 1900:
If YY > 92 and YY < 100, IYEARS = IYEARS + 99
If YY = 92 and MM > 8, IYEARS = IYEARS + 99
If YY = 92 and MM < 9, IYEARS = IYEARS + 100
If YY > 77 and YY < 92, IYEARS = IYEARS + 100

DAGE = IYEARS

EXCEPTIONS if all or part of birth date are missing:

If day of birth only is missing (DD = 00), DD = 01 was used in the BDATE formula above.

If month or year of birth are missing (MM = 00 or YY = 77), DAGE was set to age as provided on the Household Screener, if available; otherwise, DAGE was set to blanks.

DRACE (RACE/ETHNICITY F-9, F-10, AND F-11)

01 = Black not Hispanic
02 = Hispanic, Mexican
03 = Hispanic, Puerto Rican
04 = Hispanic, Cuban
05 = Hispanic, Central/South American or Other
06 = Asian
07 = American Indian/Alaskan Native
08 = Pacific Islander
09 = White
10 = Other

If background question F-10 coded 'yes' then response to background question F-11 Hispanic origin was used. If background question F-10 coded 'no' then response coded for background question F-9 was used.

DCBIRTH (COUNTRY OF BIRTH A-1)

1 = USA
2 = Spanish speaking countries
3 = European language countries
4 = Asian language countries
5 = All other countries
6 = Missing

Spanish speaking countries

01	Argentina
07	Bolivia
11	Chili
12	Columbia
13	Costa Rica
14	Cuba
17	Dominican Republic
18	Ecuador
20	El Salvador
27	Guatemala
29	Honduras
41	Mexico
44	Nicaragua
48	Panama
49	Peru
53	Puerto Rico
58	Spanish
65	Uruguay
66	Venezuela

European language countries

02	Australia
03	Austria
05	Belgium
08	Brazil
09	Bulgaria
10	Canada
15	Czechoslovakia
16	Denmark
21	England
22	Finland
23	France
24	Germany
25	Greece
31	Hungary
35	Ireland
37	Italy
42	Netherlands
43	New Zealand
46	Norway
51	Poland
52	Portugal
54	Russia
56	Scotland
59	Sweden
60	Switzerland
69	Yugoslavia

Asian language countries

30	Hong Kong
39	Japan
40	Korea
50	Philippines
61	Taiwan (China)
62	Thailand
67	Vietnam

Other

All other countries coded

DAGEARR (AGE OF ARRIVAL IN THE USA: A-2 - AGE [COMPUTED FROM F-12])

Median of A-2 [Years lived in the U.S.] - Actual Age)

A-2 1 to 5 years = 3 years
6 to 10 years = 8 years
11 to 20 years = 13 years
16 to 20 years = 18 years
21 to 30 years = 25 years
31 to 40 years = 35 years
41 to 50 years = 45 years
51 or more = 51 years

DLANGHM (LANGUAGE SPOKEN IN HOME WHILE GROWING UP A-4)

DLANGBS (LANGUAGE SPOKEN BEFORE STARTING SCHOOL A-5)

01	English only
02	English & Spanish
03	English & European
04	English & Asian
05	Spanish & Other
06	Other
07	Spanish only
08	European only
09	Asian only
10	English & Other

	"2nd Response"				
"1st Response"	None	Spanish	European	Asian	Other
English	1	2	3	4	10
Spanish	7	7	5	5	5
Other	6	5	8	9	6

English language

07	English

Spanish language

29	Spanish

European languages

04	Czech
05	Danish
06	Dutch
09	Finnish
10	Flemish
11	French
12	Gaelic
13	German
14	Greek
17	Hungarian
19	Italian
22	Latin
24	Norwegian
25	Polish
26	Portuguese
27	Romanian
28	Russian
31	Swedish
34	Ukrainian
36	Yiddish
44	Lithuanian
46	Serbo-Croatian
47	Slovak

Asian languages

03	Chinese
20	Japanese
21	Korean
32	Thai
35	Vietnamese
41	Cambodian
43	Lao, Laotian
45	Philippine

Other languages

All other language codes

DLANGRW (LANGUAGE FIRST LEARNED TO READ/WRITE A-6)

1 = English language
2 = Spanish language
3 = European languages
4 = Asian languages
5 = Other non-English languages
6 = Cannot read or write

DLANGSP (LANGUAGE USUALLY SPOKEN NOW A-15)

1 = English language
2 = Spanish language
3 = European languages
4 = Asian languages
5 = Other non-English languages

DBILING (NON-ENGLISH ORALITY A-8 PARTS A & B)

1 = Yes, non-English orality
2 = Not non-English oral

	very well or well	not well or not at all
understand	1	2
speak	1	2

DBILIT (NON-ENGLISH LITERACY A-8 PARTS C & D)

1 = Non-English literate
2 = Not non-English literate

	very well or well	not well or not at all
read	1	2
write	1	2

541

DLINGL (ENGLISH ORALITY A-17 PARTS A & B) (A-16 in the Prison sample)

1 = Yes, English orality
2 = Not English orality

	very well or well	not well or not at all
understand	1	2
speak	1	2

DLITER (ENGLISH LITERACY A-17 PARTS C & D) (A-16 in the Prison sample)

1 = Yes, English literate
2 = Not English literate

	very well or well	not well or not at all
read	1	2
write	1	2

DWEEKWG (WEEKLY WAGE D-3)

0 = weekly wage could not be computed

If D-3 per hour, multiply by number of hours indicated in D-4
If D-3 per day, multiply by number of days indicated in D-4
If D-3 per week, used amount as given
If D-3 per two-week period, divided D-3 by 2
If D-3 per month, divided D-3 by 4.3
If D-3 per year, divided D-3 by 52

DWKYRWG (AVERAGE WEEKLY WAGE D-7)

0 = weekly wage could not be computed

If D-7 per hour, multiply by number of hours indicated in D-8
If D-7 per day, multiply by number of days indicated in D-8
If D-7 per week, used amount as given
If D-7 per two-week period, divided D-8 by 2
If D-7 per month, divided D-8 by 4.3
If D-7 per year, divided D-8 by 52

DFAMINC (FAMILY INCOME F-7)

Dollar amount indicated in question F-7

01 = Under $5,000
02 = $5,000 - $9,999
03 = $10,000 - $14,999
04 = $15,000 - $19,999
05 = $20,000 - $29,999
06 = $30,000 - $39,999
07 = $40,000 - $49,999
08 = $50,000 - $74,999
09 = $75,000 - $99,999
10 = $100,000 and over
11 = Refused and I don't know

DWKSWRK (NUMBER OF WEEKS WORKED LAST YEAR D-5 PARTS A & B)

If D-5A equal to '1' weeks worked = 00

If D-5A equal to '2' weeks worked = value coded in D-5B

If D-5A equal to '3' weeks worked = 52

D-5B equal to 88 omitted or 99 incomplete, weeks worked = 00

CLFSTAT (CURRENT LABOR FORCE STATUS D-1 AND D-2)

D-1 Doing last week

Working full-time = 1
Working part-time = 2
Working two or more jobs = 3
Unemployed, laid off = 4
With job but not at work = 5
With job, family leave = 6
In school = 7
Keeping house = 8
Retired = 9
Doing volunteer work = 10
Other = 11

D-2 Have you looked for a job during past four weeks

Yes = 1
No = 2

New codes:

1 = Employed full-time (D-1 = 1 or 3)
2 = Employed part-time (D-1 = 2)
3 = Employed, not at work (D-1 = 5 or 6)
4 = Unemployed (D-1 = 4)
5 = Out of labor force (D-1 = 7, 8, 9, or 10)

ANNEARN (ANNUAL EARNINGS D-7A WEEKLY WAGE, AND DWKSWRK WEEKS WORKED)

ANNEARN = D-7A x DWKSWRK (weeks worked last year)

LABFRES (LABOR FORCE RESERVES - DERIVED FROM CLFSTAT)

0 = Not labor force reserves
1 = Labor force reserves

If age less than 65, and Current Labor Force Status equal to Unemployed
OR
with job but not at work and zero weeks worked last year

then LABFRES = 1

If age greater than or equal to 65,
OR if Current Labor Force Status is equal to 1, 2, or 3
OR if Weeks Worked is greater than 0

then LABFRES = 0

otherwise LABFRES = blank

HEDENRN (CURRENTLY ENROLLED IN HIGHER EDUCATION B-5 AND B-6)

0 = Not currently enrolled
1 = Currently enrolled

If B-5 equal '1' Yes currently enrolled and B-6 equal '2' Vocational, '3' Two year college, '4' Four or five-year college, or '5' Master's, PH.D or other advanced degree, HEDENRN = 1

If B-5 equal '2' No and B-6 equal '1' high school diploma, '6' Other or '7' None, HEDENRN = 0

BASCANY (EVER ENROLLED IN BASIC SKILLS PROGRAM B-7)

0 = No basic skills program (B-7 equal No (2))
1 = Yes basic skills program (B-7 equal Yes (1))

BASCONE (ENROLLED IN BASIC SKILLS PROGRAM IN THE LAST YEAR B-7 AND B-9)

0 = No basic skills program in the last year
1 = Yes, basic skills program within the last year

If B-7 equal '1' and B-9 A,B,C, or D = '1' or '2' BASCONE = 1
If B-7 equal '2' BASCONE = 0

BASCFIV (TAKEN BASIC SKILLS PROGRAM WITHIN THE PAST 5 YEARS B-7 AND B-9)

0 = No basic skills program within the past 5 years
1 = Yes, basic skills program within the past 5 years

If B-7 equal '1' and B-9 A,B,C, or D = '1','2' or '3' BASCONE = 1
If B-7 equal '2' BASCONE = 0

YRSLFEX (POTENTIAL YEARS OF LABOR MARKET EXPERIENCE AGE AND B-1)

YRSLFEX = Actual age - years of schooling - 5

years of schooling B-1	YRSLFEX
Still in high school	= 0
Less than high school	= grade specified, 8 if blank
Some high school	= grade specified, 11 if blank
GED/High school	= 12 years
High school graduate	= 12 years
Vocational, trade or business	= 13 years
College: less than two years	= 13 years
College: associate's degree	= 14 years
College: two years or more, no degree	= 15 years
College graduate	= 16 years
Postgraduate/No degree	= 17 years
Postgraduate/Degree	= 17 years
Don't Know	= blank
Omitted	= blank

POVLEVL (POVERTY LEVEL NUMBER OF HOUSEHOLD MEMBERS AND INCOME F-7)

0 = Not poor/near poor
1 = Poor/near poor

If number of household members is greater than zero and household income is not missing

Classified as poor/near poor

household size equal 1 and income less than $8665
household size equal 2 and income less than $11081
household size equal 3 and income less than $13575
household size equal 4 and income less than $17405
household size equal 5 and income less than $20570
household size equal 6 and income less than $23234
household size equal 7 and income less than $26322
household size equal 8 and income less than $29506
household size equal 9 and income less than $34927

All others classified as not poor/near poor

DERIVED VARIABLES FOR THE PRISON SAMPLE

DIB0801: Derived months in educational program

Derived from B-8 (months & years)
BIB0801 contains the responses to "How long ..."
BIB0802 contains the responses to "Specify ..."
but are in different units depending whether BIB0801
indicated months or years

DIB1101: Derived months in vocational program

Derived from B-11 (months & years)
BIB1101 contains the responses to "How long ..."
BIB1102 contains the responses to "Specify ..."
but are in different units depending whether BIB1101
indicated months or years

DIC0301: Derived years since arrest

Derived from C-3 (month and year arrested)
BIC0301 contains the responses to "Month"
BIC0302 contains the responses to "Year"
These two items were converted to a date value, subtracted
from the date value of the interview, and converted to
years with two decimal places.

DIC0401: Derived years since admission

Derived from C-4 (month and year of prison admission)
BIC0401 contains the responses to "Month"
BIC0402 contains the responses to "Year"

These two items were converted to a date value, subtracted from the date value of the interview, and converted to years with two decimal places.

DIC0601: Derived times on probation as a juvenile

Derived from C-6, but C-5 could be used for improvement.
BIC0501 may indicate "never" on probation, and could be used to fill in with valid data many of the blank elements that this variable now fails to count as zero.
BIC0601 contains the responses to "How many times ..."
BIC0602 contains the responses to "Specify ..."

1 = None
2 = 1 time
3 = 2 times
4 = 3-5 times
5 = 6+ times
6 = I don't know

DIC0701: Number of times on probation as an adult

Derived from C-7, but C-5 could be used for improvement.
BIC0501 may indicate "never" on probation, and could be used to fill in with valid data many of the blank elements that this variable now fails to count as zero.
BIC0701 contains the responses to "How many times ..."
BIC0702 contains the responses to "Specify ..."

1 = None
2 = 1 time
3 = 2 times
4 = 3-5 times
5 = 6+ times
6 = I don't know

DIC0901: No. of times served as a juvenile

Derived from C-9, but C-8 could be used for improvement.
BIC0801 may indicate "never" served time, and could be used to fill in with valid data many of the blank elements that this variable now fails to count as zero.
BIC0901 contains the responses to "How many times ..."
BIC0902 contains the responses to "Specify ..."

1 = None
2 = 1 time
3 = 2 times
4 = 3-5 times
5 = 6+ times
6 = I don't know

DIC1001: No. of times served as an adult

Derived from C-10, but C-8 could be used for improvement.
BIC0801 may indicate "never" served time, and could be used to

fill in with valid data many of the blank elements that this variable now fails to count as zero.
BIC1001 contains the responses to "How many times ..."
BIC1002 contains the responses to "Specify ..."

1 = None
2 = 1 time
3 = 2 times
4 = 3-5 times
5 = 6+ times
6 = I don't know

DIC1201: Derived years to release

Derived from C-12, but C-11 could be used for improvement. BIC1101 may indicate that there is no release date, and could be used to change what is now counted as valid data to a missing value, since the release date is undefined for such prisoners.
BIC1201 contains the responses to "Month"
BIC1202 contains the responses to "Year"
The date value of the interview was subtracted from the date value obtained by combining these two items, and then converting the result to years with two decimal places.

OTHER = Valid responses

DIC1401: Der. years to earliest possible release

Derived from C-14, but C-13 could be used for improvement. BIC1301 may indicate that there is no possible release date, and could be used to change what is now counted as valid data to a missing value, since the possible release date is undefined for such prisoners.
BIC1401 contains the responses to "Month"
BIC1402 contains the responses to "Year"
The date value of the interview was subtracted from the date value obtained by combining these two items, and then converting the result to years with two decimal places.

OTHER = Valid responses

DD03601: Derived hours assigned to job last week

Derived from D-3
BD03601 contains a continuous version of this item.

1 = 0-19 hours
2 = 20-29 hours
3 = 30-39 hours
4 = 40 hours
5 = 41 + hours

DID0501: Derived amount earned last month

Derived from D-5
BID0501 contains a continuous version of this item.

1 = 1 - 19 dollars
2 = 20 - 29 dollars
3 = 30 - 39 dollars
4 = 40 - 59 dollars
5 = 60 + dollars

DID1001: Derived number of months worked

Derived from D-10, but D-9 could be used for improvement.
DID0901 may indicate did not work for pay, and could be used to fill in with a valid zero many of the zero elements that this variable now considers to be missing.
BID1001 contains the responses to "How many months ..."
BID1002 contains the responses to "Specify ..."

RECODING OF DERIVED VARIABLES AND RE-RELEASE OF NEW NALS DATA FILES AND ELECTRONIC CODEBOOK

A careful review of the released dataset uncovered some minor problems with the coding of some of the derived variables. To remedy this situation, the problematic variables were recoded to reflect a more accurate definition. As a result of this recoding, the 3 data files, household, prisoner, and non-incentive, were re-issued so as to reflect these coding changes. Below we provide a list of all derived variables that were recoded in 1999 in each of the three datasets. Detailed documentation on how these variables were recoded is included in the new release of the electronic codebook (version 3) under each variable's description. Variables not rederived in version 3 maintain the same derivatives as in version 2.

Household dataset:

DAGE	derived age in years as of September 1, 1992
DRACE	derived race/ethnicity
DCBIRTH	derived country of birth
DAGEARR	age of arrival in the US
DLANGRW	derived language first learned to read/write
DBILING	derived bilingual orality
DBILIT	derived biliteracy ability
DLINGL	derived English orality (19a_b)
DLITER	derived English literacy (17c_d)
DBUSIND	derived business or industry code
DOCCUP	derived occupation codes
DWKSWRK	derived number of weeks worked last year
CLFSTAT	derived current labor force status
LABFRES	derived labor force reserves
HEDENRN	derived currently enrolled in higher education
POVLEVL	derived poverty level

Prison dataset:

DRACE	derived race/ethnicity
DCBIRTH	derived country of birth
DAGEARR	age of arrival in the US

DLANGRW	derived language first learned to read/write
DBILING	derived bilingual orality
DBILIT	derived biliteracy ability
DLINGL	derived English orality (19a_b)
DLITER	derived English literacy (17c_d)
DIB0801	derived months in educational program
DIB1101	derived months in vocational program
DIC0301	derived years since arrest
DIC0401	derived years since admission
DIC0601	derived times on probation as a juvenile
DIC0701	number of times on probation as an adult
DIC0901	number of times served as a juvenile
DIC1001	number of times served as an adult
DIC1201	derived years to release
DIC1401	derived years to earliest possible release
DID1001	derived number of months worked

Non-incentive dataset:

DAGE	derived age in years as of September 1, 1992
DRACE	derived race/ethnicity
DCBIRTH	derived country of birth
DAGEARR	age of arrival in the US
DLANGRW	derived language first learned to read/write
DBILING	derived bilingual orality
DBILIT	derived biliteracy ability
DLINGL	derived English orality (19a_b)
DLITER	derived English literacy (17c_d)
DBUSIND	derived business or industry code
DOCCUP	derived occupation codes
DWKSWRK	derived number of weeks worked last year
CLFSTAT	derived current labor force status
LABFRES	derived labor force reserves
HEDENRN	derived currently enrolled in higher education

In addition, due to a change in the definition of bilingual and biliterate ability, two new variables reflecting the new definitions were created (DTBILNG and DTBILIT) for each of the three datasets and appended at the end of each re-released dataset.

Finally, it was discovered that the program codes generated by the electronic codebook for the non-incentive sample misread the position of the second records (starting with the variable SCOR100) by 34 columns. This problem has been resolved in the new version of the electronic codebook.

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
1		ARCHITECTS/SURVEYORS
	43	Architects
	63	Surveyors & mapping scientists
2		ENGINEERS
	44	Aerospace engineers
	45	MetALLurgical+materials engineers
	46	Mining engineers
	47	Petroleum engineers
	48	Chemical engineers
	49	Nuclear engineers
	53	Civil engineers
	54	Agricultural engineers
	55	Electrical & electronic engineers
	56	Industrial engineers
	57	Mechanical engineers
	58	Marine & naval architects
	59	Engineers, n.e.c.
3	ALL	MATH/COMP SCIENTISTS
	64	Computer systems analysts & scientists
	65	Ops & systems researchers & analysts
	66	Actuaries
	67	Statisticians
	68	Mathematical scientists, n.e.c.
4	ALL	NATURAL SCIENTISTS
	69	Physicists & astronomers
	73	Chemists, excpt biochemists
	75	Geologists & geodesists
	76	Physical scientists, n.e.c.
	77	Agricultural & food scientists
	78	Biological & life scientists
	79	Forestry & conservation scientists
	83	Medical scientists
5	ALL	REGISTERED NURSES
	95	Registered nurses
6	ALL	HEALTH DIAGNOSTICS
	84	Physicians
	85	Dentists
	86	Veterinarians
	87	Optometrists
	88	Podiatrists
	89	Health diagnosing practitioners, n.e.c.

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
7	ALL	OTHER HEALTH RELATED
	96	Pharmacists
	97	Dieticians
	98	Respiratory therapists
	99	Occupational therapists
	103	Physical therapists
	104	Speech therapists
	105	Therapists, n.e.c.
	106	Physicians assistants, n.e.c.
8	ALL	ACCOUNTANTS/AUDITORS
	23	Accountants & auditors
9	ALL	PUBLIC SECT EX.& MAN
	3	Legislators
	4	Chief execs & genl admins, pub admin
	5	Admins+Offcls, Public admin
	6	Administrators, protective service
	14	Admins: education & related fields
10	ALL	PRIVAT SECT EX.& MAN
	7	Financial managers
	8	Personnel & labor relations managers
	9	Purchasing managers
	13	Managers: mkting, advertising, pub rels
	15	Managers: medicine & health
	17	Managers: food service & lodging
	18	Managers: properties & real estate
	21	Managers: service organizations, nec
	22	Managers & administrators, n.e.c.
11	ALL	OTHER MANAGEMENT
	16	Postmasters & mail superintendents
	19	Funeral directors
	24	Underwriters
	25	Other financial officers
	26	Management analysts
	27	Personnel, training & labor rels specs
	28	Purchasing agents & buyers, farm prods
	29	Buyers, wholesale & retail, excpt farm
	33	Purchasing agents & buyers, n.e.c.
	34	Business & promotion agents
	35	Construction inspectors
	36	Inspectrs+complnce offcrs,excpt const
	37	Management related occs, n.e.c.

2-DIG OCCP	**3-DIG OCCP**	**DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)**
12	ALL	TEACHERS
	113	Earth, environ, & marine sci teachers
	114	Biological science teachers
	115	Chemistry teachers
	116	Physics teachers
	117	Natural science teachers, n.e.c.
	118	Psychology teachers
	119	Economics teachers
	123	History teachers
	124	Political science teachers
	125	Sociology teachers
	126	Social science teachers, n.e.c.
	127	Engineering teachers
	128	Mathematical science teachers
	129	Computer science teachers
	133	Medical science teachers
	134	Health specialties teachers
	135	Business, commerce, & marketing teacher
	136	Agriculture & forestry teachers
	137	Art, drama & music teachers
	138	Physical education teachers
	139	Education teachers
	143	English teachers
	144	Foreign language teachers
	145	Law teachers
	146	Social work teachers
	147	Theology teachers
	148	Trade & industrial teachers
	153	Teachers, postsecondary, n.e.c.
	154	Postsecondary teachers, subj unspecifie
	155	Teachers, prekindergarten & kindergarte
	156	Teachers, elementary school
	157	Teachers, secondary school
	158	Teachers, special education
	159	Teachers, n.e.c.
13	ALL	OTHER PROFESSIONALS
	163	Counselors, educational & vocational
	164	Librarians
	165	Archivists & curators
	166	Economists
	167	Psychologists
	169	Social scientists, n.e.c.
	173	Urban planners
	174	Social workers
	175	Recreation workers
	176	Clergy
	177	Religious workers, n.e.c.

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	178	Lawyers
	179	Judges
	183	Authors
	184	Technical writers
	185	Designers
	186	Musicians & composers
	187	Actors & directors
	188	Paintrs sculptrs craftrs & art prntmkrs
	189	Photographers
	193	Dancers
	194	Artists, performers & rel wrkrs, n.e.c.
	195	Editors & reporters
	197	Public relations specialists
	198	Announcers
	199	Athletes
14	ALL	ENGINEERING TECHNIC.
	213	Electrical & electronic technicians
	214	Industrial engineering technicians
	215	Mechanical engineering technicians
	216	Engineering technicians, n.e.c.
	217	Drafting occupations
	218	Surveying & mapping technicians
15	ALL	HEALTH TECHNICIANS
	203	Clin laboratory technolgsts & technicn
	204	Dental hygienists
	205	Health record technolgsts & technicn
	206	Radiologic technicians
	207	Licensed practical nurses
	208	Health technolgsts & technicns, nec
16	ALL	SCIENCE TECHNICIANS
	223	Biological technicians
	224	Chemical technicians
	225	Science technicians, n.e.c.
17	ALL	OTHER TECHNICIANS
	226	Airplane pilots & navigators
	227	Air traffic controllers
	228	Broadcast equipment operators
	229	Computer programmers
	233	Tool programmers, numer cntl
	234	Legal assistants
	235	Technicians, n.e.c.
18	ALL	SALES REPRESENTATIVE
	253	Insurance sales occupations

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	254	Real estate sales occupations
	255	Securities & financl servs sales occups
	256	Advertising & related sales occups
	257	Sales occuptns, other business servcs
	258	Sales engineers
	259	Sales reps: mining, mfg & wholesale
19	ALL	SALES SUPERV.& PROPR
	243	Supervisors & prop, sales occupations
20	ALL	OTHER SALES RELATED
	263	Sales workers: motor vehicle & boats
	264	Sales workers: apparel
	265	Sales workers: shoes
	266	Sales workers: furniture, home furnshng
	267	Sales workers: radio, TV, hifi & applnc
	268	Sales workers: hardware, buildng supply
	269	Sales workers: parts
	274	Sales workers: other commodities
	275	Sales counter clerks
	276	Cashiers
	277	Street & door-to-door sales workers
	278	News vendors
	283	Demonstrators, promotrs & models: sales
	284	Auctioneers
	285	Sales support occupations, n.e.c.
21	ALL	ADJUSTORS AND INVEST
	375	Insurance adjustr, examiner & investgtr
	376	Investigatrs & adjustrs, exc. insurance
	377	Eligibility clerks, social welfare
	378	Bill & account collectors
22	ALL	COMPUTER EQUIP OPER.
	308	Computer operators
	309	Peripheral equipment operations
23	316	INFORMATION CLERKS
	317	Hotel clerks
	318	Transport ticket & reservation agents
	319	Receptionists
	323	Information clerks, n.e.c.
24	ALL	SECRETARIES
	313	Secretaries
25	ALL	STENOGR/TYPISTS
	314	Stenographers

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	315	Typists
26	ALL	SUPERVISORS
	303	Supervisors: general office
	304	Supervisors: computer eqpmt operators
	305	Supervisors: financial records processg
	306	Chief communications operators
	307	Supervisors: distrib, sched, adj clerks
27	ALL	OTHER ADMIN. SUPPORT
	326	Correspondence clerks
	327	Order clerks
	328	Personnel clerks, excpt payroll+timekep
	329	Library clerks
	335	File clerks
	336	Records clerks
	337	Bookkeepers, acctng & auditing clerks
	338	Payroll & timekeeping clerks
	339	Billing clerks
	343	Cost & rate clerks
	344	Billing, posting & calc machine operatr
	345	Duplicating machine operators
	346	Mail preparing & paper handling mach op
	347	Office machine operators, n.e.c.
	348	Telephone operators
	353	Communications equipment operators
	354	Postal clerks, exc. postal service
	355	Mail carriers, postal service
	356	Mail clerks, exc. postal service
	357	Messengers
	359	Dispatchers
	363	Production coordinators
	364	Traffic, shipping & receiving clerks
	365	Stock & inventory clerks
	366	Meter readers
	368	Weighers, measurers, checkers, samplers
	373	Expediters
	374	Material record, sched & distrib clerks
	379	General office clerks
	383	Bank tellers
	384	Proofreaders
	385	Data-entry keyers
	386	Statistical clerks
	387	Teachers aides
	389	Administrative support occs, n.e.c.
28	ALL	CONSTRUCTION CRAFTS
	553	Supervisors: brickmasn, stonemsn & tile

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	554	Supervisors: carpenters & reltd workers
	555	Supervrs: electrcns & powr transm instl
	556	Supervrs: paintrs, pprhangers, plasters
	557	Supervrs: plumbrs, pipefttrs, steamftrs
	558	Supervisors: construction, n.e.c.
	563	Brickmasons & stonemasons
	565	Tile setters, hard & soft
	566	Carpet instALLers
	567	Carpenters
	569	Carpenter apprentices
	573	DrywALL instALLers
	575	Electricians
	576	Electrician apprentices
	577	Electrical power instALLers & repairers
	579	Painters, construction & maintenance
	583	Paperhangers
	584	Plasterers
	585	Plumbers, pipefitters & steamfitters
	587	Plumber, pipeftr & steamftr apprentices
	588	Concrete & terrazzo finishers
	589	Glaziers
	593	Insulation workers
	594	Paving, surfacing & tamping equip opers
	595	Roofers
	596	Sheetmetal duct instALLers
	597	Structural metal workers
	598	Drillers, earth
	599	Construction trades, n.e.c.
29	ALL	OTH CRAFT/PREC.PROD
	503	Supervisors: mechanics & repairers
	505	Automobile mechanics
	507	Bus, truck, stationary engine mechanics
	508	Aircraft engine mechanics
	509	SmALL engine repairers
	514	Automobile body & releated repairers
	515	Aircraft mechanics, exc. engine
	516	Heavy equipment mechanics
	517	Farm equipment mechanics
	518	Industrial machinery repairers
	519	Machinery maintenance occupations
	523	Electronic repairers, comm & indus eqpt
	525	Data processing equipment repairers
	526	Household appliance & power tool repair
	527	Telephone line instALLers & repairers
	529	Telephone instALLers & repairers
	533	Misc electrc+electronc equipt repairers
	534	Heating, a/c, & refrigeration mechanics

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	535	Camera, watch & musical instrmnt repair
	536	Locksmith & safe repairers
	538	Office machine repairers
	539	Mechanical controls & valve repairers
	543	Elevator instALLers & repairers
	544	Millwrights
	547	Specified mechanics & repairers, n.e.c.
	549	Not specified mechanics & repairers
	613	Supervisors: extractive occupations
	614	Drillers, oil well
	615	Explosives workers
	616	Mining machine operatives
	617	Mining occupations, n.e.c.
	628	Supervisors: production occupations
	634	Tool & die makers
	635	Tool & die maker apprentices
	636	Precision assemblers, metal
	637	Machinists
	639	Machinist apprentices
	643	Boilermakers
	644	Precisn grindrs, filrs & tool sharpenrs
	645	Patternmakers & model makers, metal
	646	Lay-out workers
	647	Precious stone & metals workers (Jewel)
	649	Engravers, metal
	653	Sheet metal workers
	657	Cabinet makers & bench carpenters
	658	Furniture & wood finishers
	666	Dressmakers
	667	Tailors
	668	Upholsterers
	669	Shoe repairers
	674	Misc precision apparel & fabric workers
	675	Hand molders & shapers, exc jewelers
	677	Optical goods workers
	678	Dental lab & medical appliance technicn
	679	Bookbinders
	683	Electrical & electronic equip assembler
	684	Misc precision workers, n.e.c.
	686	Butchers & meat cutters
	687	Bakers
	688	Food batchmakers
	689	Inspectors, testers & graders
	693	Adjusters & calibrators
	694	Water & sewage treatment plant operatrs
	695	Power plant operators
	696	Stationary engineers
	699	Misc plant & system operators

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
30	ALL	TRANSPORT OPERATIVE
	803	Supervisors: motor vehicle operators
	804	Truck drivers
	806	Driver-sales workers
	808	Bus drivers
	809	Taxicab drivers & chauffeurs
	813	Parking lot attendants
	814	Motor transportation occupations, n.e.c
	823	Railroad conductors & yardmasters
	824	Locomotive operating occupations
	825	Railroad brake, signal & switch opers
	826	Rail vehicle operators, n.e.c.
	828	Ship captains & mates, exc fishng boats
	829	Sailors & deckhands
	833	Marine engineers
	843	Supervisors: material moving equipment
	844	Operating engineers
	848	Hoist & winch operators
	849	Crane & towing operators
	853	Excavating & loading machine operators
	855	Grader, dozer & scraper operators
	856	Industrial truck & tractor equip opers
	859	Misc material moving equip operators
31	ALL	FABRIC/ASSEM/INSPECT
	783	Welders & cutters
	784	Solderers & brazers
	785	Assemblers
	786	Hand cutting & trimming occupations
	787	Hand molding, casting & forming occups
	789	Hand painting, coating & decoratng occs
	793	Hand engraving & printing occupations
	795	Misc hand working occupations
	796	Productn inspectrs, checkrs & examinrs
	797	Production testers
	798	Production samplers & weighers
	799	Graders & sorters, exc agricultural
32	ALL	OTHER ASSEM/OPER/FAB
	703	Lathe & turning machine set-up operatrs
	704	Lathe & turning machine operators
	706	Punching & stamping press machine opers
	707	Rolling machine operators
	708	Drilling & boring machine operators
	709	Grindg, abradg, buffg & polishg mach op
	713	Forging machine operators
	715	Misc mtal, plastc, stone, glass work op

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	717	Fabricating machine operators, n.e.c.
	719	Molding & casting machine operators
	723	Metal plating machine operators
	724	Heat treating equipment operators
	725	Misc metal & plastic processng machn op
	726	Wood lathe, routing & planing machn ops
	727	Sawing machine operators
	728	Shaping & joining machine operators
	733	Misc woodworking machine operators
	734	Printing press operators
	735	Photoengravers & lithographers
	736	Typesetters & compositors
	737	Misc printing machine operators
	738	Winding & twisting machine operators
	739	Knittng, loopng, tapng & weavng mach op
	743	Textile cutting machine operators
	744	Textile sewing machine operators
	745	Shoe machine operators
	747	Pressing machine operators
	748	Laundering & dry cleaning machine opers
	749	Misc textile machine operators
	753	Cementing & gluing machine operators
	754	Packaging & filling machine operators
	755	Extruding & forming machine operators
	756	Mixing & blending machine operators
	757	Separatg, filterg & clarifyg mach opers
	758	Compressing & compacting mach operators
	759	Painting & paint spraying machine opers
	763	Roasting & baking machn operators, food
	764	Washing, cleaning & pickling mach opers
	765	Folding machine operators
	766	Furnace, kiln & oven operatrs, exc food
	768	Crushing & grinding machine operators
	769	Slicing & cutting machine operators
	773	Motion picture projectionists
	774	Photographic process machine operators
	777	Miscellaneous machine operatrs, n.e.c.
	779	Machine operators, not specified
33	ALL	CLEAN EQUIP.HNDL/LAB
	864	Supervrs: handlrs, eqpm cleanrs, labors
	865	Helpers: mechanics & repairers
	866	Helpers: construction trades
	869	Construction laborers
	874	Production helpers
	875	Garbage collectors
	876	Stevedores
	877	Stock handlers & baggers

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	878	Machine feeders & offbearers
	883	Freight, stock & material handlers, nec
	885	Garage & service station related occups
	887	Vehicle washers & equipment cleaners
	888	Hand packers & packagers
	889	Laborers, except construction
34	ALL	PERSONAL SERVICE OCC
	456	Supervisors: personal service occupatns
	457	Barbers
	458	Hairdressers & cosmetologists
	459	Attendants, amusement & recreatn facils
	461	Guides
	462	Ushers
	463	Public transportation attendants
	464	Baggage porters & bellhops
	465	Welfare service aides
	466	Family child care providers
	467	Early childhood teachers assistants
	468	Child care workers, n.e.c.
	469	Personal service occupations, n.e.c.
35	ALL	PUBLIC SAFETY
	413	Supervisors: fire & fire prevntn occs
	414	Supervisors: police & detectives
	415	Supervisors: guards
	416	Fire inspection & prevention occs
	417	Firefighting occupations
	418	Police & detectives, public servant
	423	Sheriffs, bailiffs & othr law enforcmnt
	424	Correctional institution officers
36	ALL	HEALTH SERVICES
	445	Dental assistants
	446	Health aides, exc. nursing
	447	Nursing aides, orderlies & attendants
37	ALL	OTHER SERVICES
	404	Cooks, private household
	405	Housekeepers & butlers
	406	Child care workers, private household
	407	Private household cleaners & servants
	425	Crossing guards
	426	Guards & police, exc. public service
	427	Protective service occs, n.e.c.
	433	Supervisors: food prep & service occpns
	434	Bartenders
	435	Waiters & waitresses

2-DIG OCCP	3-DIG OCCP	DERIVED OCCUPATION CODES (DOCCUP) & Detailed Census occupations (BLD1001)
	436	Cooks
	438	Food counter, fountain & related occpns
	439	Kitchen workers, food preparation
	443	Assistants to waiters/waitresses
	444	Miscellaneous food preparation occupns
	448	Supervisors: cleaning & bldg servc
	449	Maids & housemen
	453	Janitors & cleaners
	454	Elevator operators
	455	Pest control occupations
38	ALL	MANAGER/OPERATORS
	473	Farmers, exc. horticultural
	474	Horticultural specialty farmers
	475	Managers: farms, exc. horticultural
	476	Managers: horticultural specialty farms
	494	Supervisors: forestry & logging workers
	497	Captains & othr officers, fishing vessl
39	ALL	OTHER FARM/FISH/HUNT
	477	Supervisors: farm workers
	479	Farm workers
	484	Nursery workers
	485	Supervisors: related agricultural occs
	486	Groundskeepers & gardeners, exc. farm
	487	Animal caretakers, except farm
	488	Graders & sorters, agricultural prods
	489	Inspectors, agricultural products
	495	Forestry workers, except logging
	496	Timber cutting & logging occupations
	498	Fishers
40	ALL	MILITARY
	903	Commissioned officers & warrant officrs
	904	Non-commissioned offcrs & othr enlisted
	905	Military occupation, rank not specified
41	ALL	{MISSING}
	999	{DOES NOT APPLY}
42	ALL	NOT WORKED PAST 3 YR
	999	{DOES NOT APPLY}
43	ALL	NEVER WORKED

Appendix J

SPECIAL CODES FOR CONTINUOUS VARIABLES

Appendix J: Special Codes for Continuous Variables
NATIONAL ADULT LITERACY SURVEY

J.1 INTRODUCTION

This appendix lists special codes and their meanings for continuous variables and noncognitive items in Blocks ID, WT and BQ. Continuous variables and items that are not shown here do not have special codes associated with them. For discrete variables and items, labels for all values can be found in the data set codebooks. The codes for background questionnaire item BLA0102, "Country of Birth," are provided in Appendix K; that item was designated continuous because the discrete numerical data included a large number of valid responses which could not be listed in the codebooks.

Special codes that apply to variables and items in all samples are listed in Section J.2; special codes for variables and items which are found only in the Prison sample are presented in Section J.3.

J.2 SPECIAL CODES FOR CONTINUOUS FIELDS APPEARING IN ALL SAMPLES

Field Name	Special Code	Description
BQTIME	999	Missing
EXTIME	999	Missing
DAGEARR	0	Missing
DWEEKWG	0	Missing
DWKYRWG	0	Missing
BLA0102	0	Missing, Skipped In Error
	99	Missing
BLD0301	77777	Respondent Refused
	88888	Missing
	99999	Could Not Be Computed
BLD0801	77777	Respondent Refused
	88888	Missing
	99999	Could Not Be Computed
BLD0901	999	Missing
BLD1001	999	Missing
BD07502	77	Year of Birth Omitted

*Note: If the entire date of birth question had no response, Month of Birth (BD07501), Day of Birth (BLF1201) and BD07502 are **all** coded "00."*

J.3 SPECIAL CODES FOR CONTINUOUS FIELDS THAT APPEAR ONLY IN THE PRISON SAMPLE

Field Name	Special Code	Description
DIB0801	0	Missing
DIB1101	0	Missing
DID1001	0	Missing
BIB0802	99	Missing
BIC0302	0	Missing
BIC0402	0	Missing
BIC0602	0	Missing
BIC0902	0	Missing
BIC1002	0	Missing
BIC1202	1991	Missing
BIC1402	1991	Missing
BD03601	888	Don't Know, Not Specified
	999	Missing, Skipped In Error
BID0501	77777	Respondent Refused
	88888	Missing
	99999	Could Not Be Computed
BID1002	888	Don't Know
	999	Missing

Appendix K

COUNTRY OF BIRTH CODES

Appendix K: Country of Birth Codes
NATIONAL ADULT LITERACY SURVEY

Table K-1 provides information about the valid data values and their associated labels for background questionnaire item A-1, Country of Birth (field name BLA0102).

Table K-1. Country codes

01 Argentina
02 Australia
03 Austria
04 Bahamas
05 Belgium
06 Bermuda
07 Bolivia
08 Brazil
09 Bulgaria
10 Canada
11 Chile
12 Colombia
13 Costa Rica
14 Cuba
15 Czechoslovakia
16 Denmark
17 Dominican Rep.
18 Ecuador
19 Egypt
20 El Salvador
21 England
22 Finland
23 France
24 Germany
25 Greece
26 Guam
27 Guatemala
28 Haiti
29 Honduras
30 Hong Kong
31 Hungary
32 India
33 Indonesia
34 Iran
35 Ireland
36 Israel
37 Italy
38 Jamaica
39 Japan
40 Korea
41 Mexico
42 Netherlands
43 New Zealand
44 Nicaragua
45 Nigeria
46 Norway
47 Pakistan
48 Panama
49 Peru
50 Philippines
51 Poland
52 Portugal
53 Puerto Rico
54 Russia
55 Saudi Arabia
56 Scotland
57 South Africa
58 Spain
59 Sweden
60 Switzerland
61 Taiwan (China)
62 Thailand
63 Turkey
64 Utd. Arab Emir.
65 Uruguay
66 Venezuela
67 Vietnam
68 Virgin Islands
69 Yugoslavia
70 United States

71 Other

00 Skipped in Error *[No territory or country code was specified in BLA0102, although BLA0101 was coded "2" (U.S. Territory) or "3" (Other Country).]*

Appendix L

NOTES ON SCORING

Appendix L: Notes on Scoring
NATIONAL ADULT LITERACY SURVEY

L.1 INTRODUCTION

This appendix provides information about scoring. Four topics are covered: 1) the scoring for five cognitive items for which scoring keys could not be provided; 2) the reasons four cognitive items were omitted from scaling; 3) the method by which items were scored Right, Wrong, Omitted and Not Reached for input to the computer programs used for scaling; and 4) a table of exceptions to the general item scoring method, with information about the method applied to each exceptional item.

L.2 SCORING FOR CERTAIN COGNITIVE ITEMS

For five cognitive items (AB60502, AB60503, AB70801, AB70802, and N120801), scoring keys are not provided, as the scoring could not be expressed as a correct value or range of values. Items AB60502 and AB60503 were scored together as a single item; AB70801 and AB70802 were also scored jointly. Scoring for N120801 simply could not be expressed as a range of values. Details on the scoring of each of these items are presented below. Refer to the codebooks for information about data values which occurred for the individual items.

AB60502 and AB60503

These items were scored jointly to produce a single score for input to the computer programs used for scaling. The combined scoring method was as follows:

If the respondent scored "**1**" on **either or both** items, the score was **Right**;
a "**2**" on **both** items was **Wrong**.

Note: For all respondents, scores of "I Don't Know" (7), "Omitted" (8), and "Not Reached" (9), when given, applied to both items. See Section L.4 for scoring information about these categories.

AB70801 and AB70802

These items were scored jointly to produce a single score for input to the computer programs used for scaling. The combined scoring method was as follows:

If the respondent scored "**1**," "**2**," "**3**," "**4**," "**5**," or "**6**" on AB70801, the score was **Right**;
if not, a score of "**1**," **2**," "**3**," "**4**," "**5**," or "**6**" on AB70802 was **Wrong**;
otherwise, a "**0**" on **both** was **Omit**; also,
an "**8**" (Newspaper Missing) on **both** was **Not Reached**,
and a "**99**" **on AB70801** was **Not Reached**.

Note: Respondents with "I Don't Know" (77) or "Newspaper Missing" (8) scores received them for both items. The general scoring method for "I Don't Know," described in Section L.4, applies to this item.

N120801

The values "1" and "3" are the correct keys for this item.

L.3 COGNITIVE ITEMS OMITTED FROM SCALING

IRT scaling parameters are not given for four cognitive items, AB40904, AB60503, AB70105 and AB70802. The reasons these items were omitted from scaling are provided below.

AB40904	This item was considered to measure a secondary trait, because a correct score on this item was dependent upon having received a correct score to the previous question, AB40901.
AB60503	This item was scored jointly with AB60502 to produce a single score for input to the scaling programs. The IRT parameters for the joint "item" are listed with item AB60502.
AB70105	Due to a poor fit with the IRT model, this item was excluded from calibration.
AB70802	This item was scored jointly with AB70801 to produce a single score for input to the scaling programs. The IRT parameters for the joint "item" are listed with item AB70801.

L.4 ITEM SCORING FOR SCALING PURPOSES -- GENERAL METHOD

For input to the computer programs used for IRT scaling, all possible item responses are assigned to one of four categories: Wrong (0), Right (1), Omitted (2), or Not Reached (3). This section explains the method applied to most of the NALS cognitive items to convert individual item data values to these four categories. Section L.5 contains information about exceptions to this general method.

Refer to the codebooks for the data values possible for each individual cognitive item.

The correct key(s) for the item were considered **Right** (1).

Nonresponses *which are followed by valid responses to other items in the same block (section)* were considered **Omitted** (2). In the public-use files and in the codebooks, these nonresponses are designated by the special codes for "Omitted."

Nonresponses *which occur* ***after*** *the last item in a block (section) with a valid response* were considered **Not Reached** (3). In the public-use files and in the codebooks, these nonresponses are designated by the special codes for "Not Reached."

Multiple responses were considered **Wrong** (0).

"I Don't Know" responses were considered **Wrong** (0).

All other responses were considered **Wrong** (0).

L.5 ITEM SCORING FOR SCALING PURPOSES -- SPECIAL CASES

For some items, the scoring for scaling purposes was more complex than the method described in Section L.4. In some cases, an item's potential data values could include values with meanings such as "Illegible," "Off Task," "Incomprehensible," or "Newspaper Missing." The numerical value associated with these categories varied from item to item; as an aid to the user, the associated scoring category for each of these data values is listed individually for each item in Table L-1.

In other cases, items which had been administered in earlier literacy surveys were scored for scaling as they had been scaled previously, in order to make trend comparisons possible. All items that had to be scored by a method other than that described in Section L.4 are listed in Table L-1. The table shows to which of the four scoring categories each valid response to an item would be assigned. **Boldface** item responses in Table L-1 may be considered exceptions to the general method.

Table L-1. Exceptions to general scoring method for scaling

Item Field Name	Right (1)	Wrong (0)	Omitted (2)	Not Reached (3)
AB21101	1	2, 7	8	**3**, 9
AB21201	1	2, 7	8	**3**, 9
AB30501	1	2, 7	8	**3**, 9
AB30601	2, 3, 4	1, **7**	**77**, 88	99
AB31201	2, 3	1, **7**, **8**	**77**, 88	99
AB40901	1	2, **6**, 7	8	**3**, 9
AB40904	0, 1, 2, 3, 4, 5	**6**	8	9
AB50101	1	2, 7	8	**3**, 9
AB50301	2, 3	1, **7**	**77**, 88	99
AB60201	2, 3, 4	1, **7**, **8**	**77**, 88	99
AB60502, AB60503	*These items were combined for scoring. See Section L.2 for details.*			
AB60601	1	2, 7	8	**3**, 9
AB70104	1	2	**7**, 8	9
AB70105	1	2	**7**, 8	9
AB70801, AB70802	*These items were combined for scoring. See Section L.2 for details.*			

Table L-1. Exceptions to general scoring method for scaling — Continued

Item Field Name	Right (1)	Wrong (0)	Omitted (2)	Not Reached (3)
AB71201	1	2, 7	8	3, 9
N081001	1, 2	3, 4, 7	8	5, 9
N090601	1	2, 7	8	3, 9
N090701	1	2, 7	8	3, 9
N090801	1	2, 7	8	3, 9
N090901	1	2, 3, 4, 5, 7	8	6, 9
N091001	1	2, 3, 4, 5, 6, 7, 77	88	8, 99
N110302	1	2, 3, 7	4, 8	9
N110303	1, 2	3, 4, 7	5, 8	9
N120901	1	2, 7	8	3, 9
N121001	1	2, 7	8	3, 9
N121101	1, 2	3, 7	8	4, 9
N130101	1	2, 3, 8, 77	4, 88	99
N130102	1	2, 8, 77	3, 88	99
N130103	1	2, 3, 8, 77	4, 88	99
N130104	1	2, 3, 8, 77	4, 88	99

Appendix M

NOTES ON VARIABLES

Appendix M: Notes on Variables
NATIONAL ADULT LITERACY SURVEY

M.1 INTRODUCTION

This appendix provides sample-specific information for the users of certain variables. It is divided into sections for each sample.

M.2 MAIN (HOUSEHOLD) SAMPLE

The Taylor Series Linearization Variance Stratum and Paired Sampling Unit identifiers (TSLVSTR and TSLVPSU) can be used to estimate standard errors using a Taylor Series Linearization package such as SUDAAN. **CAUTION:** When using SUDAAN to produce standard errors for an individual state, *the user should process the entire file and specify the state in a SUBPOPN statement*, rather than processing a separate analysis file containing records for that state alone. This procedure is necessary because in some instances the National PSU sample design disregarded state boundaries in its stratification scheme.

M.3 PRISON SAMPLE

For the Taylor Series Linearization Variance Paired Sampling Unit identifier (TSLVPSU), "3" is a valid number for this sample.

M.4 NON-INCENTIVE SAMPLE

Some continuous variables in Block WT of the Non-incentive sample are provided in a different format than variables with the same names that appear in other samples. The user should note that the Person Base Weight (BASEWT), Step 1 and 2 Raking Factors (S1RAKFC and S2RAKFC), Final Weight (WEIGHT), and the sixty replicate weights (REPL01 - REPL60) have a field width of 11 in the Non-incentive sample data file. For other samples, these variables have a field width of 10. The number of implicit decimal places (4) is unchanged, however.

Appendix N

IMPACT OF TREATMENT ON DISTRIBUTION OF SCALE SCORES

Appendix N: Impact of Treatment on Distribution of Scale Scores
NATIONAL ADULT LITERACY SURVEY

by Andrew Kolstad, National Center for Education Statistics

Combined with other background information, there is strong evidence to support the notion that nonresponse to the cognitive tasks was not a random occurrence. Procedures to incorporate proficiency values for nonresponding individuals were established based on the field study and were thus in place before the National Adult Literacy Survey data analyses were conducted. The selected procedure used self-reported reasons for nonresponse to assign wrong answers to all the missing cognitive responses of adults when the reasons were related to literacy skills. If the reasons were unrelated to literacy, then no wrong answers were assigned. After the wrong answers were (or were not) assigned to the cognitive tasks that were presented each adult, normal scaling procedures for creating plausible values were applied. The normal procedure uses statistical imputation to estimate plausible values, even in the absence of any cognitive data.

Table N-1 shows the average proficiency scores resulting from this procedure, based on the full-scale survey. Adults who completed the assessment had average proficiencies ranging from 279 to 285 on the three literacy scales. Because the missing responses of adults who were unable to complete the assessment for reasons related to literacy were treated as wrong answers, the average proficiency scores of these adults were considerably lower, ranging from 114 to 124. The proficiency scores of nearly all who were unable to complete the assessment were in the Level 1 range (below 225). Adults who stopped for other reasons or for unstated reasons had scores between those of the other two groups, ranging from 228 to 237. These adults were not found only in the lowest literacy level, but were distributed across the five levels.

Those who were assigned wrong answers did not score zero on the three scales. Zero is the score value given to wrong answers, but the resulting scale scores of those given wrong answers were from 114 to 124, depending on the scale. Zero has no intrinsic meaning on the adult literacy scales, since each was derived from a latent variable with a mean of zero and a variance of one, transformed by adding a constant and multiplying by a fixed factor in order to eliminate negative numbers and decimal points in the resulting scale.

Table N-1. Percentage and average proficiencies on each scale, by assessment completion status: Adults, 1992

Assessment completion status		Average proficiency		
	Percent	Prose literacy	Document literacy	Quantitative literacy
Total	100%	272	267	271
Completed assessment	88%	285	279	284
Did not completed assessment for literacy-related reasons	6%	124	116	114
Did not complete assessment for reasons unrelated to literacy	6%	237	228	231

N.1 The distribution of performance

The method chosen for placing those with incomplete cognitive data on the literacy scale takes many factors into account, and as a result does not put those with missing data into a single point along the scales. The distributions on each scale are shown in Exhibits N-1, N-2, and N-3. These distributions were generated from a 3000-case subsample of the 1992 data, selected with a probability inversely proportional to the case weight, thereby providing a self-weighting subsample. The method of sample selection also took into account the segment from which the respondents were selected, so that the average number of adults selected from each segment fell from 6.9 to 1.3, considerably reducing any possible impact of clustering on the findings. The distributions were generated using a nonparametric method for estimating a density function, and can be thought of as a smoothed histogram (SAS Institute, 1991). In effect, the fixed intervals of the histogram are replaced by a moving window of selected width, within which each observation is smoothed. Two densities are plotted for each scale: one for the complete data, and one for the data leaving out any cases with less than sufficient cognitive data (for any reason).

The area between the two curves represents the locations along the distributions where the partially or totally missing data appear. It is reassuring to note that the distributions are unimodal, and that the methods selected for treating the missing data did not assign all such cases to the same point along the literacy scales, and very few, if any, cases occurred at the zero point on the scales. Nevertheless, most of the cases with missing data occur in the bottom half of the distribution, so leaving out the cases with missing cognitive data would clearly have biased the findings of the survey in an upward direction.

Exhibit N-1. Prose literacy distribution of U.S. adults, with and without missing cognitive data

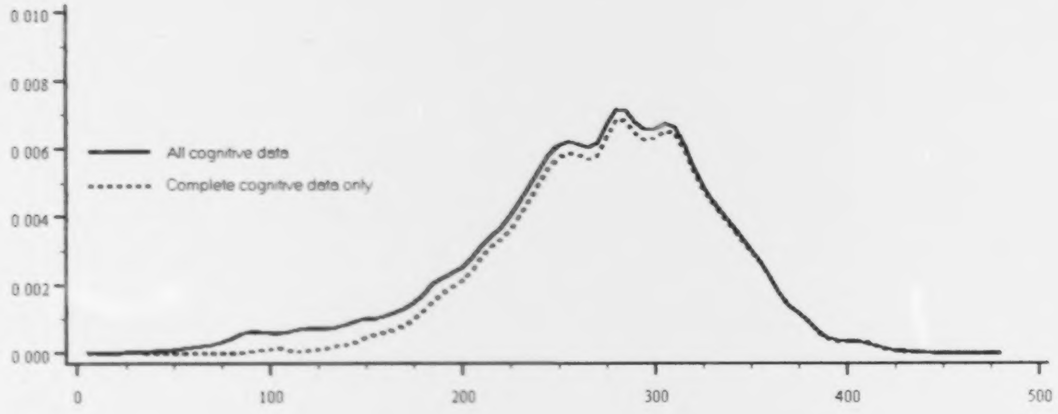


Exhibit N-2. Document literacy distribution of U.S. adults, with and without missing cognitive data

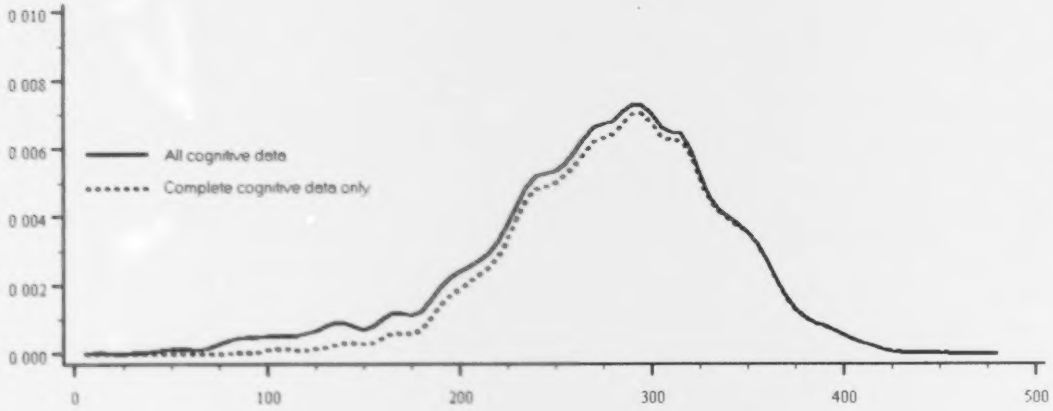
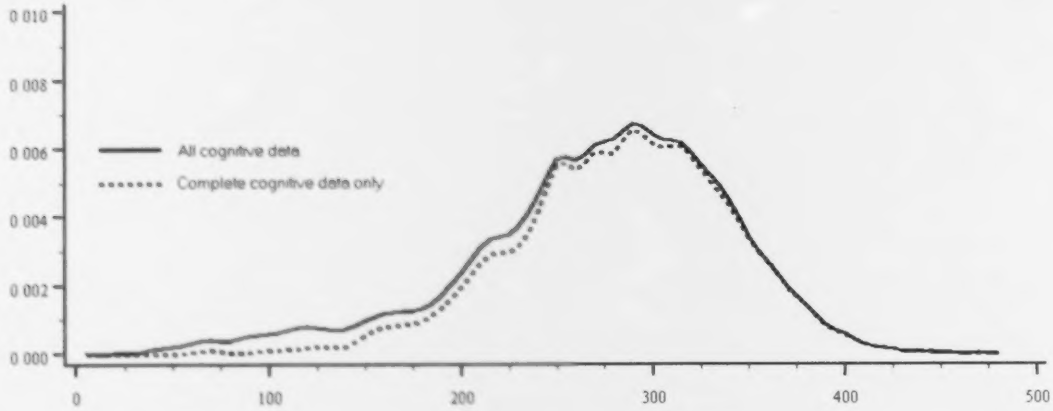


Exhibit N-3. Quantitative literacy distribution of U.S. adults, with and without missing cognitive data



BEST COPY AVAILABLE

N.2 Performance in the lowest literacy level

Nearly all of those who were assigned wrong answers to tasks that had been left blank scored below 225 on the three scales (Level 1, the lowest of the five literacy levels). Some of those in Level 1 appear there through their own efforts, while others were helped to be there by the method used for treating missing cognitive data.

Table N-2 presents the percentages of adults who performed at least one task correctly and their average proficiencies on each scale. Most adults in Level 1 performed at least one literacy task correctly. On the prose literacy scale, 72 percent of adults in Level 1 performed at least one task correctly, as did 83 percent on the document scale and 66 percent on the quantitative scale. A small proportion of adults in Level 1 failed to perform correctly any literacy task. This group constituted 26 percent of adults on the prose literacy scale, 11 percent of adults on the document literacy scale, and 21 percent of adults on the quantitative literacy scale. Wide disparities in average proficiencies occurred between those who performed at least one task correctly and those who did not. The average proficiencies of these adults (94 on the document scale, 110 on the prose scale, and 113 on the quantitative scale) were considerably lower than those of adults who did succeed with at least one task, whose scores were 182 on the document scale and 190 on the prose and quantitative scales.

Some adults in Level 1 (6 to 8 percent) did not respond to any of the literacy tasks for reasons unrelated to their literacy skills or for unknown reasons. These persons could not be described as either meeting or failing to meet the demands of literacy tasks, so they are distinguished as a separate group. Their proficiencies were inferred from the performance of other adults with similar demographic backgrounds and fell in the middle range between the other two groups.

Table N-2. Percentage who performed at least one task correctly and average proficiency on each scale: Adults in Level 1, 1992

Performance on literacy tasks	Prose literacy		Document literacy		Quantitative literacy	
	Percent	Average proficiency	Percent	Average proficiency	Percent	Average proficiency
Total in Level 1	100%	173	100%	172	100%	167
At least one literacy task correct	72%	190	83%	182	66%	190
No literacy tasks Correct	21%	113	11%	94	26%	110
No literacy task data	7%	177	6%	177	8%	159

Nearly all adults who correctly responded to at least one literacy task also completed the assessment. Still, some adults broke off the assessment after already having shown some initial success. Table 8-8 divides adults in Level 1 into four groups by crossing two variables: those who did not with who were successful with at least one task and those who were not successful with any task, crossed with those who completed the assessment (at least five literacy tasks) and those who did not.

Table N-3. Percentage who performed at least one task or no tasks correctly and average proficiency on each scale by provision of sufficient cognitive data: Adults in Level 1, 1992

Completion of literacy tasks	Prose literacy		Document literacy		Quantitative literacy	
	Percent	Average proficiency	Percent	Average proficiency	Percent	Average proficiency
Total in Level 1 with at least one literacy task correct	100%	190	100%	182	100%	190
Provided sufficient cognitive data	87%	196	83%	192	90%	194
Did not provided sufficient cognitive data	13%	153	17%	132	10%	153
Total in Level 1 with no literacy tasks correct	100%	113	100%	94	100%	110
Provided sufficient cognitive data	14%	148	2%	-	14%	146
Did not provided sufficient cognitive data	86%	107	98%	93	86%	98

Across the scales, from 83 to 90 percent of those in Level 1 who correctly responded to at least one task also completed the assessment. Their average scores ranged from 192 to 196. The remainder (10 to 17 percent) performed at least one task correctly before breaking off the assessment. Their average scores were much lower, ranging from 132 to 153.

The population of adults who scored in Level 1 on each scale includes not only those who demonstrated the skills needed to succeed with at least some of the tasks in Level 1—who constituted the majority—but also those who were unable to succeed with any of the tasks in this level. Nearly all of those in Level 1 who failed to succeed with at least one literacy task also failed to complete the assessment, as shown in the following table. Most of these adults (from 86 to 98 percent) either did not start or broke off the assessment for literacy-related reasons, so that any literacy tasks that remained unanswered were treated as wrong answers. Their scores averaged from 93 to 107 across the scales.

Some respondents completed the assessment, yet failed to respond correctly to any literacy tasks (2 to 14 percent of those in Level 1). These individuals had scores that averaged 148 on the prose scale and 146 on the quantitative scale; too few cases were available to estimate an average score on the document scale.

The pattern of Level 1 proficiencies associated with various combinations of missing and incorrect answers show the consequences of including, rather than leaving out adults who were unable to complete the assessment for literacy-related reasons. In general, the very low scores of these individuals bring down the average for any group in which they are a significant component. Omitting these individuals from the assessment would have resulted in inflated estimates of the literacy skills of the adult population overall and particularly of certain populations.

Appendix O

ESTIMATED COMPOSITE FACTORS FOR A SELECTED SET OF DEMOGRAPHIC VARIABLES FOR ILLINOIS, INDIANA, LOUISIANA, NEW JERSEY, NEW YORK, OHIO, PENNSYLVANIA, TEXAS, AND WASHINGTON

Appendix O: Estimated Composite Factors For a Selected Set of Demographic Variables for Illinois, Indiana, Louisiana, New Jersey, New York, Ohio, Pennsylvania, Texas, and Washington

NATIONAL ADULT LITERACY SURVEY

Table O-1. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Illinois sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.5464	0.6411	0.3993	0.5710	0.7771	0.4620
	no	0.5441	0.5579	0.3956	0.5155	0.7481	0.4690
Gender							
Male	yes	0.5261	0.5355	0.4338	0.5785	0.7692	0.4419
	no	0.5115	0.5159	0.4316	0.5576	0.7818	0.4458
Female	yes	0.6939	0.5584	0.6607	0.7406	0.6359	0.5915
	no	0.6275	0.3878	0.6077	0.6847	0.5412	0.4836
Race/Ethnicity							
Black	yes	0.4681	0.4501	0.4996	0.2987	0.7064	0.2533
	no	0.5954	0.5715	0.6300	0.4091	0.7983	0.3666
Hispanic	yes	0.4572	0.5878	0.5665	0.8150	0.5676	0.4273
	no	0.1358	0.1829	0.1987	0.4118	0.1764	0.1222
Other	yes	0.7161	0.6489	0.4957	0.5865	0.8113	0.5974
	no	0.6445	0.5015	0.4087	0.4650	0.7299	0.5223
Education Level							
No HS degree	yes	0.6695	0.7686	0.7239	0.5581	0.6011	0.4673
	no	0.5657	0.6972	0.6725	0.4672	0.5066	0.3837
HS degree	yes	0.5728	0.4787	0.5841	0.6271	0.7432	0.5029
	no	0.5975	0.4390	0.5307	0.6061	0.7279	0.5495
Some college	yes	0.5436	0.6278	0.4127	0.6293	0.8168	0.5884
	no	0.5377	0.6090	0.4694	0.5945	0.7805	0.5699
College graduate	yes	0.4919	0.6718	0.6155	0.4793	0.6922	0.6104
	no	0.4320	0.6593	0.5531	0.4265	0.6665	0.5648
Age							
16 to 24 years	yes	0.5398	0.5100	0.6016	0.8138	0.5723	0.5422
	no	0.4249	0.3597	0.5184	0.7235	0.5088	0.5064
25 to 44 years	yes	0.4551	0.6861	0.4475	0.5318	0.7820	0.4593
	no	0.4627	0.6339	0.4448	0.4912	0.7700	0.4340
45 to 64 years	yes	0.5845	0.6562	0.6628	0.6718	0.6148	0.5796
	no	0.5641	0.5676	0.5574	0.5960	0.5273	0.5444
Country of Birth							
Not USA	yes	0.3770	0.5388	0.6225	0.8245	0.4339	0.3319
	no	0.1234	0.1997	0.2817	0.5057	0.1519	0.1078
USA	yes	0.6014	0.7062	0.4441	0.5519	0.7610	0.5469
	no	0.6227	0.6476	0.4457	0.5087	0.7354	0.5047

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-2. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Illinois sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.4290	0.6861	0.6988	0.8946	0.8489	0.5364
	no	0.3930	0.6775	0.6683	0.8411	0.8175	0.5031
Gender							
Male	yes	0.5108	0.5965	0.6610	0.7680	0.6388	0.5300
	no	0.4477	0.5748	0.6226	0.7210	0.6251	0.5036
Female	yes	0.4633	0.7341	0.5312	0.7538	0.7811	0.5746
	no	0.4358	0.6702	0.5041	0.6844	0.7068	0.4570
Race/Ethnicity							
Black	yes	0.3925	0.5264	0.5123	0.4193	0.3105	0.2740
	no	0.5125	0.6423	0.6454	0.5375	0.4184	0.3890
Hispanic	yes	0.4332	0.5660	0.5604	0.6526	0.7372	0.3840
	no	0.1177	0.1792	0.1786	0.2362	0.3070	0.1072
Other	yes	0.5904	0.6874	0.7313	0.8563	0.8586	0.6751
	no	0.5253	0.6485	0.6618	0.7461	0.7929	0.5804
Education Level							
No HS degree	yes	0.5791	0.6635	0.7074	0.6076	0.5059	0.5137
	no	0.4608	0.5516	0.6135	0.5152	0.4102	0.4185
HS degree	yes	0.4638	0.5379	0.6224	0.6427	0.6378	0.5199
	no	0.4687	0.5422	0.5991	0.6314	0.6538	0.5602
Some college	yes	0.4313	0.7843	0.8861	0.7278	0.7936	0.5991
	no	0.3863	0.7491	0.8399	0.6776	0.7000	0.5407
College graduate	yes	0.5485	0.6465	0.4289	0.7870	0.7431	0.6471
	no	0.4989	0.5839	0.4275	0.7673	0.7392	0.6328
Age							
16 to 24 years	yes	0.6070	0.8713	0.8184	0.6641	0.6938	0.5346
	no	0.4895	0.7474	0.7523	0.5420	0.5964	0.4744
25 to 44 years	yes	0.2719	0.4667	0.5446	0.8482	0.8142	0.5297
	no	0.3039	0.5065	0.5647	0.8179	0.8091	0.4918
45 to 64 years	yes	0.6949	0.6495	0.6570	0.7280	0.8006	0.6816
	no	0.6155	0.6073	0.6083	0.6745	0.7514	0.6210
Country of Birth							
Not USA	yes	0.4573	0.5083	0.6151	0.7494	0.6692	0.4120
	no	0.1584	0.1772	0.2613	0.4115	0.3233	0.1445
USA	yes	0.4000	0.6094	0.6749	0.8565	0.8401	0.6156
	no	0.3925	0.6291	0.6510	0.7948	0.8088	0.5433

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-3. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Illinois sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.3068	0.5788	0.6494	0.7088	0.7533	0.4386
	no	0.3402	0.5334	0.5642	0.6810	0.7282	0.4256
Gender							
Male	yes	0.4756	0.4434	0.6978	0.6905	0.6715	0.4060
	no	0.4563	0.4009	0.6575	0.6783	0.6880	0.3696
Female	yes	0.4973	0.8474	0.7255	0.7160	0.6886	0.5707
	no	0.4623	0.7830	0.6506	0.6620	0.5725	0.4676
Race/Ethnicity							
Black	yes	0.4519	0.4857	0.4256	0.4028	0.6921	0.2449
	no	0.5867	0.6074	0.5510	0.5248	0.7862	0.3645
Hispanic	yes	0.3776	0.6876	0.5199	0.6559	0.1936	0.3911
	no	0.0999	0.2711	0.1564	0.2465	0.0315	0.1083
Other	yes	0.5894	0.6199	0.6598	0.6981	0.7824	0.6761
	no	0.5670	0.5329	0.5089	0.6180	0.6927	0.5547
Education Level							
No HS degree	yes	0.3703	0.6443	0.4686	0.5407	0.4986	0.3966
	no	0.3197	0.4763	0.3997	0.4836	0.3850	0.3539
HS degree	yes	0.4596	0.5509	0.5104	0.5099	0.8620	0.4106
	no	0.5175	0.5921	0.4646	0.4932	0.8059	0.4728
Some college	yes	0.5083	0.4545	0.7222	0.8476	0.7199	0.5858
	no	0.4923	0.4340	0.6725	0.8190	0.6872	0.5129
College graduate	yes	0.3866	0.4967	0.5502	0.4051	0.6392	0.7269
	no	0.3426	0.4376	0.4768	0.4046	0.5911	0.6694
Age							
16 to 24 years	yes	0.5433	0.6593	0.5262	0.4458	0.6118	0.5689
	no	0.4865	0.5335	0.4297	0.3586	0.5525	0.5415
25 to 44 years	yes	0.3086	0.6166	0.6452	0.7891	0.7788	0.3889
	no	0.3678	0.6058	0.6162	0.7679	0.7651	0.3758
45 to 64 years	yes	0.6311	0.5077	0.6454	0.7771	0.6874	0.6445
	no	0.5629	0.4353	0.5420	0.7383	0.5751	0.5891
Country of Birth							
Not USA	yes	0.3971	0.3981	0.5244	0.8076	0.6418	0.4070
	no	0.1318	0.1287	0.1920	0.4857	0.2838	0.1389
USA	yes	0.4226	0.5962	0.7033	0.7182	0.7680	0.4995
	no	0.4307	0.5681	0.6425	0.7100	0.7516	0.4883

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-4. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Indiana sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.6940	0.7012	0.5557	0.6744	0.8524	0.6322
Gender							
Male	yes	–	–	–	–	–	–
	no	0.6771	0.6950	0.6075	0.7229	0.8837	0.6137
Female	yes	–	–	–	–	–	–
	no	0.7480	0.5117	0.7403	0.8028	0.6783	0.6137
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.4634	0.4109	0.5432	0.2536	0.6827	0.2934
Hispanic	yes	–	–	–	–	–	–
	no	0.5928	0.7109	0.6971	0.8840	0.6900	0.5570
Other	yes	–	–	–	–	–	–
	no	0.7876	0.6771	0.5855	0.6474	0.8506	0.6942
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.8130	0.8827	0.8768	0.7461	0.7772	0.6798
HS degree	yes	–	–	–	–	–	–
	no	0.7610	0.6123	0.6950	0.7583	0.8487	0.7261
Some college	yes	–	–	–	–	–	–
	no	0.6494	0.7067	0.5867	0.6958	0.8425	0.6724
College graduate	yes	–	–	–	–	–	–
	no	0.4980	0.7193	0.6183	0.4934	0.7245	0.6300
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.6082	0.5229	0.6907	0.8404	0.6834	0.6892
25 to 44 years	yes	–	–	–	–	–	–
	no	0.6231	0.7689	0.6088	0.6566	0.8674	0.5924
45 to 64 years	yes	–	–	–	–	–	–
	no	0.7053	0.6991	0.6924	0.7289	0.6663	0.6895
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.5734	0.7172	0.7872	0.9105	0.6328	0.5260
USA	yes	–	–	–	–	–	–
	no	0.7443	0.7527	0.5851	0.6400	0.8265	0.6319

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-5. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Indiana sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.5384	0.8045	0.7939	0.9030	0.8962	0.6551
Gender							
Male	yes	–	–	–	–	–	–
	no	0.6115	0.7301	0.7609	0.8314	0.7750	0.6641
Female	yes	–	–	–	–	–	–
	no	0.5815	0.7806	0.6580	0.7966	0.8127	0.5886
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.3814	0.4867	0.5539	0.3711	0.2455	0.3043
Hispanic	yes	–	–	–	–	–	–
	no	0.5589	0.6889	0.6809	0.7624	0.8258	0.5111
Other	yes	–	–	–	–	–	–
	no	0.6856	0.7942	0.8061	0.8502	0.8913	0.7329
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.7218	0.8048	0.8416	0.7714	0.7023	0.6980
HS degree	yes	–	–	–	–	–	–
	no	0.6467	0.7198	0.7649	0.7719	0.8126	0.7243
Some college	yes	–	–	–	–	–	–
	no	0.4928	0.8200	0.8853	0.7600	0.7729	0.6409
College graduate	yes	–	–	–	–	–	–
	no	0.5663	0.6479	0.4982	0.8130	0.7897	0.6955
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.6635	0.8484	0.8615	0.7028	0.7524	0.6625
25 to 44 years	yes	–	–	–	–	–	–
	no	0.4533	0.6724	0.7288	0.8937	0.8939	0.6397
45 to 64 years	yes	–	–	–	–	–	–
	no	0.7372	0.7445	0.7389	0.7873	0.8478	0.7473
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.6375	0.6824	0.7748	0.8708	0.8196	0.6076
USA	yes	–	–	–	–	–	–
	no	0.5171	0.7528	0.7603	0.8586	0.8773	0.6598

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-6. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Indiana sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.4853	0.6825	0.7090	0.8089	0.8413	0.5861
Gender							
Male	yes	–	–	–	–	–	–
	no	0.6147	0.5648	0.7963	0.8104	0.8235	0.5314
Female	yes	–	–	–	–	–	–
	no	0.6066	0.8639	0.7726	0.7861	0.7000	0.6008
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.5121	0.4725	0.4001	0.3724	0.6666	0.3146
Hispanic	yes	–	–	–	–	–	–
	no	0.5091	0.7889	0.6530	0.7716	0.2896	0.5265
Other	yes	–	–	–	–	–	–
	no	0.7276	0.7042	0.6883	0.7777	0.8264	0.7135
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.6010	0.7368	0.6896	0.7803	0.6915	0.6490
HS degree	yes	–	–	–	–	–	–
	no	0.6929	0.7626	0.6312	0.6752	0.8803	0.6648
Some college	Yes	–	–	–	–	–	–
	No	0.6031	0.5432	0.7565	0.8733	0.7707	0.6130
College graduate	yes	–	–	–	–	–	–
	no	0.4063	0.5047	0.5421	0.4749	0.6548	0.7258
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.6646	0.6882	0.6066	0.5409	0.7175	0.7183
25 to 44 years	yes	–	–	–	–	–	–
	no	0.5318	0.7488	0.7531	0.8647	0.8637	0.5294
45 to 64 years	yes	–	–	–	–	–	–
	no	0.6949	0.5840	0.6794	0.8380	0.7092	0.7158
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.5884	0.5894	0.6987	0.9008	0.7892	0.6002
USA	yes	–	–	–	–	–	–
	no	0.5606	0.6888	0.7513	0.8143	0.8413	0.6255

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-7. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Louisiana sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU*	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.8509	0.8339	0.7579	0.8200	0.9297	0.8134
Gender							
Male	yes	–	–	–	–	–	–
	no	0.7799	0.7688	0.7237	0.8021	0.9274	0.7426
Female	yes	–	–	–	–	–	–
	no	0.8961	0.7318	0.8945	0.9227	0.8556	0.8168
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.7650	0.6631	0.8451	0.4387	0.8711	0.6559
Hispanic	yes	–	–	–	–	–	–
	no	0.8623	0.9149	0.9110	0.9710	0.9037	0.8379
Other	yes	–	–	–	–	–	–
	no	0.9165	0.8611	0.8070	0.8442	0.9438	0.8702
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.8874	0.9369	0.9427	0.8533	0.8712	0.8122
HS degree	yes	–	–	–	–	–	–
	no	0.9006	0.8151	0.8633	0.8981	0.9401	0.8838
Some college	yes	–	–	–	–	–	–
	no	0.7933	0.8382	0.7839	0.8134	0.9167	0.8150
College graduate	yes	–	–	–	–	–	–
	no	0.7260	0.8731	0.8117	0.7200	0.8738	0.8185
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.7597	0.6908	0.8313	0.9159	0.8325	0.8430
25 to 44 years	yes	–	–	–	–	–	–
	no	0.8191	0.8895	0.8070	0.8257	0.9447	0.7906
45 to 64 years	yes	–	–	–	–	–	–
	no	0.8413	0.8154	0.7964	0.8336	0.7898	0.8241
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.9337	0.9637	0.9749	0.9906	0.9480	0.9204
USA	yes	–	–	–	–	–	–
	no	0.8793	0.8642	0.7706	0.7902	0.9129	0.7908

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-8. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Louisiana sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.7403	0.9095	0.9003	0.9525	0.9503	0.8192
Gender							
Male	yes	–	–	–	–	–	–
	no	0.7141	0.8176	0.8439	0.8910	0.8457	0.7734
Female	yes	–	–	–	–	–	–
	no	0.8099	0.9116	0.8566	0.9179	0.9241	0.7985
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.6798	0.7307	0.8648	0.6046	0.3454	0.6547
Hispanic	yes	–	–	–	–	–	–
	no	0.8362	0.9032	0.8975	0.9320	0.9522	0.8096
Other	yes	–	–	–	–	–	–
	no	0.8664	0.9192	0.9245	0.9443	0.9602	0.8908
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.8231	0.8782	0.9079	0.8683	0.8199	0.8175
HS degree	yes	–	–	–	–	–	–
	no	0.8385	0.8785	0.9002	0.9057	0.9241	0.8833
Some college	yes	–	–	–	–	–	–
	no	0.6478	0.9003	0.9366	0.8639	0.8544	0.7761
College graduate	yes	–	–	–	–	–	–
	no	0.7759	0.8323	0.7282	0.9205	0.9084	0.8604
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.8026	0.9160	0.9308	0.8312	0.8655	0.8171
25 to 44 years	yes	–	–	–	–	–	–
	no	0.7043	0.8554	0.8824	0.9542	0.9571	0.8208
45 to 64 years	yes	–	–	–	–	–	–
	no	0.8473	0.8539	0.8536	0.8838	0.9136	0.8554
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.9464	0.9562	0.9726	0.9862	0.9795	0.9412
USA	yes	–	–	–	–	–	–
	no	0.7261	0.8823	0.8785	0.9273	0.9375	0.8041

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-9. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Louisiana sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	**National Certainty PSU?**	**State Compositing Factor: Beta**					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.7198	0.8330	0.8360	0.9068	0.9243	0.7767
Gender							
Male	yes	–	–	–	–	–	–
	no	0.7456	0.6845	0.8535	0.8761	0.8876	0.6587
Female	yes	–	–	–	–	–	–
	no	0.8242	0.9464	0.9064	0.9157	0.8637	0.8091
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.8343	0.7555	0.6802	0.6427	0.8613	0.7152
Hispanic	yes	–	–	–	–	–	–
	no	0.8142	0.9402	0.8921	0.9391	0.6370	0.8273
Other	yes	–	–	–	–	–	–
	no	0.8876	0.8755	0.8669	0.9114	0.9335	0.8809
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.7752	0.8038	0.8286	0.9002	0.7974	0.8195
HS degree	yes	–	–	–	–	–	–
	no	0.8671	0.9017	0.8272	0.8531	0.9538	0.8511
Some college	yes	–	–	–	–	–	–
	no	0.7601	0.7163	0.8599	0.9316	0.8737	0.7484
College graduate	yes	–	–	–	–	–	–
	no	0.6456	0.7305	0.7537	0.7118	0.8348	0.8754
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.8223	0.8257	0.7700	0.7171	0.8557	0.8630
25 to 44 years	yes	–	–	–	–	–	–
	no	0.7722	0.8875	0.8868	0.9423	0.9429	0.7498
45 to 64 years	yes	–	–	–	–	–	–
	no	0.8234	0.7304	0.7880	0.9133	0.7983	0.8437
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.9364	0.9370	0.9594	0.9894	0.9742	0.9393
USA	yes	–	–	–	–	–	–
	no	0.7626	0.8358	0.8579	0.9077	0.9214	0.7934

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-10. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in New Jersey sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.5931	0.6837	0.4458	0.6169	0.8084	0.5096
	no	0.6327	0.6783	0.4834	0.6073	0.8094	0.5455
Gender							
Male	yes	0.5628	0.5720	0.4705	0.6141	0.7944	0.4786
	no	0.5967	0.5703	0.5023	0.6267	0.8094	0.5318
Female	yes	0.7388	0.6121	0.7084	0.7808	0.6855	0.6437
	no	0.7502	0.5933	0.7107	0.7693	0.6728	0.6555
Race/Ethnicity							
Black	yes	0.6267	0.6096	0.6557	0.4483	0.8210	0.3928
	no	0.5188	0.4792	0.6025	0.3076	0.7351	0.3505
Hispanic	yes	0.4471	0.5778	0.5565	0.8088	0.5576	0.4174
	no	0.5868	0.6718	0.6872	0.8644	0.6650	0.5635
Other	yes	0.6825	0.6117	0.4558	0.5472	0.7856	0.5583
	no	0.7382	0.6464	0.5219	0.5887	0.8145	0.6210
Education Level							
No HS degree	yes	0.6797	0.7768	0.7331	0.5695	0.6122	0.4789
	no	0.7318	0.8249	0.7898	0.6370	0.6724	0.5483
HS degree	yes	0.7533	0.6765	0.7618	0.7929	0.8682	0.6973
	no	0.6848	0.5705	0.6611	0.7166	0.8080	0.6329
Some college	yes	0.4308	0.5174	0.3087	0.5190	0.7391	0.4760
	no	0.4942	0.6015	0.4308	0.5692	0.7979	0.5599
College graduate	yes	0.5539	0.7242	0.6724	0.5414	0.7426	0.6677
	no	0.5685	0.7457	0.6862	0.5458	0.7432	0.6753
Age							
16 to 24 years	yes	0.5702	0.5407	0.6307	0.8317	0.6021	0.5725
	no	0.5544	0.5418	0.6311	0.8297	0.6087	0.5780
25 to 44 years	yes	0.5530	0.7640	0.5453	0.6271	0.8416	0.5571
	no	0.5777	0.7408	0.5547	0.5959	0.8383	0.5612
45 to 64 years	yes	0.5415	0.6158	0.6227	0.6321	0.5726	0.5365
	no	0.5813	0.6234	0.6106	0.6306	0.5763	0.5612
Country of Birth							
Not USA	yes	0.4392	0.6019	0.6809	0.8588	0.4979	0.3913
	no	0.4867	0.6274	0.7247	0.8743	0.5441	0.4503
USA	yes	0.6365	0.7361	0.4811	0.5884	0.7870	0.5834
	no	0.6676	0.7311	0.4989	0.5777	0.7866	0.5875

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-11. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in New Jersey sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	**National Certainty PSU?**	**State Compositing Factor: Beta**					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	0.4762	0.7256	0.7374	0.9113	0.8718	0.5834
	no	0.5256	0.7413	0.7482	0.9145	0.8692	0.6086
Gender							
Male	yes	0.5477	0.6315	0.6934	0.7933	0.6723	0.5667
	no	0.5747	0.6599	0.7284	0.8180	0.6877	0.6097
Female	yes	0.5186	0.7750	0.5857	0.7925	0.8166	0.6276
	no	0.5600	0.7870	0.5985	0.7854	0.8052	0.6311
Race/Ethnicity							
Black	yes	0.5521	0.6795	0.6670	0.5793	0.4620	0.4186
	no	0.4607	0.5635	0.5915	0.4458	0.3073	0.3662
Hispanic	yes	0.4232	0.5560	0.5503	0.6433	0.7292	0.3744
	no	0.5575	0.6687	0.6707	0.7399	0.8051	0.5280
Other	yes	0.5512	0.6521	0.6987	0.8355	0.8380	0.6391
	no	0.6371	0.7167	0.7426	0.8655	0.8587	0.7032
Education Level							
No HS degree	yes	0.5904	0.6738	0.7169	0.6186	0.5175	0.5254
	no	0.6656	0.7230	0.7661	0.6916	0.5824	0.6016
HS degree	yes	0.6633	0.7261	0.7896	0.8038	0.8004	0.7114
	no	0.5795	0.6348	0.6917	0.7463	0.7189	0.6614
Some college	yes	0.3252	0.6979	0.8317	0.6295	0.7096	0.4870
	no	0.3727	0.7515	0.8673	0.6960	0.7472	0.5510
College graduate	yes	0.6091	0.7011	0.4907	0.8258	0.7877	0.7016
	no	0.6191	0.7288	0.5255	0.8345	0.7912	0.7305
Age							
16 to 24 years	yes	0.6359	0.8845	0.8360	0.6910	0.7193	0.5650
	no	0.6274	0.8844	0.8360	0.6905	0.7154	0.5583
25 to 44 years	yes	0.3561	0.5644	0.6391	0.8922	0.8665	0.6252
	no	0.4117	0.5951	0.6367	0.8872	0.8630	0.6385
45 to 64 years	yes	0.6566	0.6087	0.6165	0.6920	0.7712	0.6425
	no	0.6744	0.6170	0.6378	0.7140	0.7714	0.6580
Country of Birth							
Not USA	yes	0.5216	0.5723	0.6741	0.7946	0.7236	0.4756
	no	0.5721	0.6013	0.7084	0.8229	0.7618	0.5361
USA	yes	0.4362	0.6442	0.7067	0.8738	0.8591	0.6502
	no	0.4882	0.6635	0.7169	0.8771	0.8543	0.6508

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-12. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in New Jersey sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	**National Certainty PSU?**	**State Compositing Factor: Beta**					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	0.3488	0.6244	0.6915	0.7465	0.7870	0.4860
	no	0.4410	0.6392	0.6784	0.7443	0.7859	0.5164
Gender							
Male	yes	0.5126	0.4802	0.7281	0.7212	0.7033	0.4421
	no	0.5757	0.5184	0.7349	0.7434	0.7239	0.4844
Female	yes	0.5525	0.8739	0.7673	0.7588	0.7339	0.6238
	no	0.5891	0.8684	0.7543	0.7496	0.7177	0.6330
Race/Ethnicity							
Black	yes	0.6112	0.6430	0.5856	0.5627	0.8108	0.3821
	no	0.5733	0.5460	0.4558	0.4367	0.7244	0.3580
Hispanic	yes	0.3680	0.6787	0.5098	0.6467	0.1873	0.3814
	no	0.5036	0.7746	0.6257	0.7432	0.2348	0.5228
Other	yes	0.5502	0.5815	0.6230	0.6634	0.7540	0.6401
	no	0.6398	0.6368	0.6495	0.7036	0.7854	0.6934
Education Level							
No HS degree	yes	0.3812	0.6549	0.4802	0.5522	0.5103	0.4078
	no	0.4759	0.7033	0.5614	0.6140	0.5563	0.4982
HS degree	yes	0.6594	0.7363	0.7036	0.7031	0.9343	0.6133
	no	0.6115	0.6680	0.6058	0.5978	0.9066	0.5504
Some college	yes	0.3964	0.3461	0.6229	0.7795	0.6202	0.4733
	no	0.4663	0.4223	0.6856	0.8222	0.6894	0.5315
College graduate	yes	0.4470	0.5587	0.6108	0.4662	0.6944	0.7734
	no	0.4624	0.5733	0.5918	0.5197	0.7083	0.7821
Age							
16 to 24 years	yes	0.5736	0.6864	0.5568	0.4764	0.6406	0.5988
	no	0.5818	0.6984	0.5577	0.4684	0.6526	0.6079
25 to 44 years	yes	0.3980	0.7043	0.7292	0.8471	0.8391	0.4852
	no	0.4637	0.7069	0.7277	0.8414	0.8361	0.5082
45 to 64 years	yes	0.5895	0.4640	0.6044	0.7454	0.6487	0.6034
	no	0.6159	0.4704	0.5937	0.7589	0.6243	0.6436
Country of Birth							
Not USA	yes	0.4601	0.4612	0.5879	0.8445	0.6986	0.4704
	no	0.5109	0.5021	0.6239	0.8663	0.7371	0.5253
USA	yes	0.4593	0.6315	0.7334	0.7473	0.7934	0.5367
	no	0.5127	0.6485	0.7171	0.7423	0.7902	0.5437

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-13. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in New York sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	0.3839	0.4802	0.2559	0.4077	0.6433	0.3075
	no	0.5603	0.5563	0.4107	0.5158	0.7517	0.4855
Gender							
Male	yes	0.4079	0.4171	0.3223	0.4600	0.6741	0.3295
	no	0.5318	0.5305	0.4524	0.5741	0.7965	0.4698
Female	yes	0.5032	0.3610	0.4653	0.5605	0.4383	0.3928
	no	0.6298	0.3728	0.6088	0.6799	0.5333	0.4826
Race/Ethnicity							
Black	yes	0.3206	0.3051	0.3487	0.1860	0.5633	0.1539
	no	0.1762	0.1465	0.2447	0.0739	0.3459	0.1041
Hispanic	yes	0.2520	0.3632	0.3433	0.6379	0.3443	0.2299
	no	0.5215	0.6463	0.6348	0.8501	0.6218	0.4824
Other	yes	0.6385	0.5641	0.4077	0.4983	0.7507	0.5096
	no	0.6775	0.5290	0.4442	0.4958	0.7543	0.5594
Education Level							
No HS degree	yes	0.4431	0.5662	0.5074	0.3316	0.3719	0.2563
	no	0.6748	0.7839	0.7652	0.5817	0.6208	0.4977
HS degree	yes	0.4285	0.3393	0.4399	0.4847	0.6181	0.3614
	no	0.6589	0.4917	0.5804	0.6585	0.7710	0.6170
Some college	yes	0.3471	0.4296	0.2387	0.4311	0.6656	0.3896
	no	0.4613	0.5478	0.4321	0.5088	0.7326	0.5067
College graduate	yes	0.3993	0.5844	0.5236	0.3873	0.6070	0.5183
	no	0.3747	0.6127	0.4940	0.3642	0.6053	0.5027
Age							
16 to 24 years	yes	0.3136	0.2884	0.3704	0.6299	0.3426	0.3156
	no	0.4211	0.3619	0.5186	0.7238	0.5102	0.5065
25 to 44 years	yes	0.3403	0.5743	0.3333	0.4122	0.6889	0.3440
	no	0.4647	0.6145	0.4420	0.4717	0.7639	0.4277
45 to 64 years	yes	0.3624	0.4354	0.4427	0.4527	0.3920	0.3578
	no	0.6028	0.5898	0.5736	0.6187	0.5501	0.5798
Country of Birth							
Not USA	yes	0.2299	0.3656	0.4486	0.6986	0.2744	0.1968
	no	0.5602	0.7035	0.7791	0.9047	0.6220	0.5132
USA	yes	0.4651	0.5807	0.3153	0.4152	0.6472	0.4102
	no	0.6110	0.6138	0.4259	0.4719	0.7107	0.4753

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-14. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in New York sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.2798	0.5306	0.5454	0.8145	0.7440	0.3744
	no	0.4068	0.6879	0.6754	0.8410	0.8183	0.5137
Gender							
Male	yes	0.3933	0.4785	0.5476	0.6727	0.5233	0.4117
	no	0.4630	0.5935	0.6405	0.7347	0.6419	0.5242
Female	yes	0.2783	0.5522	0.3361	0.5776	0.6146	0.3764
	no	0.4483	0.6728	0.5111	0.6796	0.6990	0.4528
Race/Ethnicity							
Black	yes	0.2574	0.3734	0.3603	0.2791	0.1945	0.1683
	no	0.1442	0.1949	0.2359	0.1287	0.0704	0.1089
Hispanic	yes	0.2341	0.3429	0.3377	0.4290	0.5287	0.1996
	no	0.4820	0.6210	0.6110	0.7041	0.7773	0.4366
Other	yes	0.5023	0.6063	0.6559	0.8067	0.8096	0.5927
	no	0.5645	0.6851	0.6933	0.7671	0.8127	0.6146
Education Level							
No HS degree	yes	0.3508	0.4365	0.4871	0.3782	0.2869	0.2933
	no	0.5701	0.6628	0.7164	0.6242	0.5257	0.5303
HS degree	yes	0.3261	0.3943	0.4797	0.5015	0.4962	0.3771
	no	0.5300	0.6036	0.6530	0.6828	0.7119	0.6241
Some college	yes	0.2529	0.6187	0.7764	0.5441	0.6319	0.4001
	no	0.3086	0.6886	0.7959	0.6132	0.6184	0.4633
College graduate	yes	0.4548	0.5567	0.3403	0.7174	0.6652	0.5574
	no	0.4377	0.5348	0.3832	0.7237	0.6888	0.5886
Age							
16 to 24 years	yes	0.3756	0.7250	0.6371	0.4351	0.4688	0.3091
	no	0.4879	0.7497	0.7527	0.5432	0.5958	0.4717
25 to 44 years	yes	0.1873	0.3507	0.4247	0.7752	0.7301	0.4102
	no	0.3139	0.5133	0.5614	0.8094	0.8041	0.4870
45 to 64 years	yes	0.4792	0.4282	0.4363	0.5196	0.6186	0.4638
	no	0.6391	0.6380	0.6385	0.7004	0.7705	0.6477
Country of Birth							
Not USA	yes	0.2936	0.3378	0.4408	0.5960	0.4995	0.2569
	no	0.6167	0.6623	0.7621	0.8658	0.8125	0.5937
USA	yes	0.2775	0.4734	0.5447	0.7747	0.7517	0.4800
	no	0.3840	0.6130	0.6284	0.7731	0.7849	0.5101

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-15. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in New York sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.1863	0.4154	0.4893	0.5572	0.6123	0.2878
	no	0.3649	0.5407	0.5598	0.6850	0.7321	0.4388
Gender							
Male	yes	0.3602	0.3309	0.5890	0.5807	0.5592	0.2979
	no	0.4796	0.4187	0.6699	0.6952	0.7069	0.3878
Female	yes	0.3066	0.7127	0.5414	0.5298	0.4969	0.3726
	no	0.4734	0.7774	0.6433	0.6587	0.5602	0.4662
Race/Ethnicity							
Black	yes	0.3066	0.3362	0.2844	0.2657	0.5465	0.1481
	no	0.2267	0.1912	0.1375	0.1280	0.3338	0.1118
Hispanic	yes	0.1953	0.4682	0.3023	0.4327	0.0876	0.2044
	no	0.4358	0.7344	0.5846	0.7181	0.2303	0.4547
Other	yes	0.5013	0.5332	0.5760	0.6183	0.7158	0.5937
	no	0.6094	0.5682	0.5348	0.6503	0.7190	0.5862
Education Level							
No HS degree	yes	0.1876	0.4157	0.2573	0.3162	0.2810	0.2052
	no	0.4217	0.5893	0.5119	0.6032	0.5048	0.4623
HS degree	yes	0.3223	0.4069	0.3683	0.3678	0.7774	0.2804
	no	0.5862	0.6574	0.5161	0.5516	0.8307	0.5454
Some college	yes	0.3157	0.2711	0.5371	0.7129	0.5343	0.3870
	no	0.4213	0.3708	0.6045	0.7701	0.6253	0.4314
College graduate	yes	0.3021	0.4040	0.4566	0.3186	0.5489	0.6464
	no	0.2919	0.3801	0.3960	0.3690	0.5339	0.6130
Age							
16 to 24 years	yes	0.3166	0.4298	0.3019	0.2386	0.3803	0.3394
	no	0.4880	0.5386	0.4305	0.3566	0.5554	0.5424
25 to 44 years	yes	0.2161	0.4982	0.5289	0.6979	0.6849	0.2821
	no	0.3825	0.6007	0.6080	0.7598	0.7589	0.3753
45 to 64 years	yes	0.4087	0.2941	0.4238	0.5849	0.4705	0.4228
	no	0.5912	0.4616	0.5592	0.7620	0.5867	0.6201
Country of Birth							
Not USA	yes	0.2452	0.2460	0.3523	0.6743	0.4692	0.2530
	no	0.5720	0.5730	0.6805	0.8945	0.7742	0.5850
USA	yes	0.2966	0.4597	0.5773	0.5949	0.6561	0.3651
	no	0.4234	0.5461	0.6011	0.6831	0.7269	0.4628

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-16. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Ohio sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	**National Certainty PSU?**	**State Compositing Factor: Beta**					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.5993	0.5867	0.4530	0.5725	0.7909	0.5396
Gender							
Male	yes	–	–	–	–	–	–
	no	0.5716	0.6046	0.5032	0.6289	0.8411	0.5043
Female	yes	–	–	–	–	–	–
	no	0.6469	0.3681	0.6498	0.7287	0.5652	0.4827
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.3799	0.3208	0.4602	0.1859	0.5974	0.2283
Hispanic	yes	–	–	–	–	–	–
	no	0.7310	0.8264	0.8146	0.9366	0.8068	0.6941
Other	yes	–	–	–	–	–	–
	no	0.7028	0.5585	0.4724	0.5353	0.7821	0.5960
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.7032	0.8062	0.8064	0.6209	0.6603	0.5451
HS degree	yes	–	–	–	–	–	–
	no	0.6486	0.4559	0.5433	0.6298	0.7547	0.6120
Some college	yes	–	–	–	–	–	–
	no	0.5886	0.6268	0.5295	0.6208	0.7739	0.5899
College graduate	yes	–	–	–	–	–	–
	no	0.4911	0.7260	0.6084	0.4972	0.7414	0.6337
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.5336	0.4193	0.6264	0.7850	0.6255	0.6529
25 to 44 years	yes	–	–	–	–	–	–
	no	0.5086	0.6631	0.4940	0.5408	0.8029	0.4673
45 to 64 years	yes	–	–	–	–	–	–
	no	0.6216	0.5902	0.5813	0.6343	0.5566	0.6030
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.4291	0.5869	0.6740	0.8507	0.4905	0.3827
USA	yes	–	–	–	–	–	–
	no	0.6665	0.6499	0.4883	0.5339	0.7547	0.5194

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-17. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Ohio sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.4168	0.7361	0.7135	0.8415	0.8468	0.5471
Gender							
Male	yes	–	–	–	–	–	–
	no	0.4841	0.6308	0.6579	0.7444	0.6925	0.5494
Female	yes	–	–	–	–	–	–
	no	0.4725	0.6855	0.5683	0.7107	0.7292	0.4555
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.2940	0.3881	0.4824	0.2815	0.1741	0.2354
Hispanic	yes	–	–	–	–	–	–
	no	0.6947	0.8054	0.7958	0.8586	0.8991	0.6505
Other	yes	–	–	–	–	–	–
	no	0.5784	0.7231	0.7323	0.7636	0.8396	0.6271
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.5776	0.6904	0.7446	0.6483	0.5674	0.5584
HS degree	yes	–	–	–	–	–	–
	no	0.5063	0.5948	0.6432	0.6472	0.7217	0.6020
Some college	yes	–	–	–	–	–	–
	no	0.4017	0.7523	0.8220	0.6703	0.6641	0.5381
College graduate	yes	–	–	–	–	–	–
	no	0.5672	0.6263	0.5084	0.8187	0.8121	0.6963
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.5853	0.7761	0.8192	0.6194	0.6880	0.6193
25 to 44 years	yes	–	–	–	–	–	–
	no	0.3447	0.5781	0.6475	0.8308	0.8429	0.5094
45 to 64 years	yes	–	–	–	–	–	–
	no	0.6336	0.6687	0.6512	0.7029	0.7859	0.6561
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.4967	0.5468	0.6584	0.7902	0.7174	0.4640
USA	yes	–	–	–	–	–	–
	no	0.4010	0.6827	0.6703	0.7751	0.8178	0.5368

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-18. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Ohio sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.3800	0.5756	0.5973	0.7403	0.7807	0.4824
Gender							
Male	yes	–	–	–	–	–	–
	no	0.4933	0.4401	0.7096	0.7333	0.7645	0.4079
Female	yes	–	–	–	–	–	–
	no	0.4968	0.7931	0.6827	0.7089	0.5803	0.4718
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.4305	0.3797	0.3178	0.2886	0.5774	0.2547
Hispanic	yes	–	–	–	–	–	–
	no	0.6571	0.8745	0.7832	0.8695	0.4429	0.6748
Other	yes	–	–	–	–	–	–
	no	0.6430	0.6052	0.5732	0.6988	0.7502	0.5996
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.4622	0.5853	0.5587	0.6861	0.5539	0.5216
HS degree	yes	–	–	–	–	–	–
	no	0.5710	0.6579	0.4732	0.5356	0.7842	0.5472
Some college	yes	–	–	–	–	–	–
	no	0.5256	0.4552	0.6682	0.8204	0.6934	0.5020
College graduate	yes	–	–	–	–	–	–
	no	0.4047	0.4973	0.5536	0.4706	0.6510	0.7177
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.6085	0.5946	0.5291	0.4729	0.6595	0.6859
25 to 44 years	yes	–	–	–	–	–	–
	no	0.4297	0.6496	0.6472	0.7956	0.7968	0.4041
45 to 64 years	yes	–	–	–	–	–	–
	no	0.5890	0.4767	0.5677	0.7747	0.6004	0.6135
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.4445	0.4456	0.5653	0.8357	0.6775	0.4565
USA	yes	–	–	–	–	–	–
	no	0.4503	0.5831	0.6516	0.7533	0.7803	0.5285

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-19. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Pennsylvania sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.5376	0.6328	0.3908	0.5622	0.7709	0.4532
	no	0.6075	0.6120	0.4605	0.5842	0.7970	0.5408
Gender							
Male	yes	0.5142	0.5236	0.4222	0.5668	0.7606	0.4301
	no	0.5795	0.6010	0.5069	0.6315	0.8375	0.5132
Female	yes	0.6860	0.5493	0.6524	0.7334	0.6273	0.5825
	no	0.6708	0.4207	0.6619	0.7371	0.5927	0.5222
Race/Ethnicity							
Black	yes	0.4661	0.4481	0.4976	0.2970	0.7047	0.2517
	no	0.5759	0.5464	0.6143	0.3843	0.7828	0.3512
Hispanic	yes	0.0939	0.1493	0.1386	0.3516	0.1391	0.0841
	no	0.8437	0.9025	0.8946	0.9663	0.8934	[illegible]241
Other	yes	0.6926	0.6228	0.4675	0.5589	0.7934	0.5699
	no	0.6649	0.5087	0.4292	0.4815	0.7438	0.5486
Education Level							
No HS degree	yes	0.5977	0.7090	0.6579	0.4808	0.5250	0.3914
	no	0.7040	0.8049	0.7924	0.6154	0.6545	0.5348
HS degree	yes	0.5925	0.4989	0.6036	0.6458	0.7583	0.5231
	no	0.6572	0.4791	0.5677	0.6487	0.7667	0.6176
Some college	yes	0.5293	0.6143	0.3988	0.6158	0.8080	0.5744
	no	0.5924	0.6485	0.5223	0.6407	0.8016	0.6115
College graduate	yes	0.4753	0.6571	0.5997	0.4628	0.6779	0.5946
	no	0.4716	0.7028	0.5909	0.4715	0.7132	0.6093
Age							
16 to 24 years	yes	0.5359	0.5061	0.5979	0.8114	0.5684	0.5383
	no	0.4326	0.3611	0.5257	0.7269	0.5168	0.5177
25 to 44 years	yes	0.4875	0.7133	0.4797	0.5639	0.8033	0.4916
	no	0.5382	0.7018	0.5238	0.5760	0.8224	0.5059
45 to 64 years	yes	0.4828	0.5588	0.5661	0.5759	0.5143	0.4778
	no	0.6565	0.6391	0.6322	0.6771	0.6054	0.6396
Country of Birth							
Not USA	yes	0.2677	0.4136	0.4989	0.7394	0.3164	0.2308
	no	0.1820	0.2938	0.3796	0.6271	0.2206	0.1565
USA	yes	0.5939	0.6997	0.4365	0.5442	0.7553	0.5391
	no	0.6739	0.6800	0.5000	0.5562	0.7705	0.5463

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-20. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Pennsylvania sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.4203	0.6784	0.6913	0.8912	0.8443	0.5276
	no	0.4407	0.7379	0.7235	0.8612	0.8540	0.5630
Gender							
Male	yes	0.4989	0.5849	0.6503	0.7594	0.6278	0.5181
	no	0.5010	0.6390	0.6725	0.7589	0.6945	0.5628
Female	yes	0.4541	0.7268	0.5220	0.7468	0.7747	0.5655
	no	0.4868	0.7094	0.5683	0.7297	0.7497	0.4963
Race/Ethnicity							
Black	yes	0.3906	0.5243	0.5102	0.4173	0.3087	0.2724
	no	0.4868	0.6171	0.6334	0.5098	0.3901	0.3712
Hispanic	yes	0.0860	0.1383	0.1356	0.1878	0.2566	0.0713
	no	0.8266	0.8927	0.8892	0.9234	0.9471	0.7954
Other	yes	0.5628	0.6627	0.7085	0.8418	0.8443	0.6498
	no	0.5461	0.6804	0.6866	0.7423	0.8059	0.5939
Education Level							
No HS degree	yes	0.5022	0.5912	0.6394	0.5317	0.4288	0.4366
	no	0.5925	0.6929	0.7437	0.6508	0.5617	0.5601
HS degree	yes	0.4840	0.5580	0.6412	0.6611	0.6563	0.5400
	no	0.5223	0.6036	0.6533	0.6688	0.7201	0.6161
Some college	yes	0.4173	0.7744	0.8801	0.7163	0.7840	0.5851
	no	0.4295	0.7771	0.8527	0.7060	0.7194	0.5768
College graduate	yes	0.5321	0.6312	0.4128	0.7757	0.7303	0.6318
	no	0.5433	0.6156	0.4788	0.8008	0.7842	0.6752
Age							
16 to 24 years	yes	0.6032	0.8695	0.8161	0.6606	0.6904	0.5306
	no	0.4956	0.7462	0.7569	0.5457	0.6025	0.4858
25 to 44 years	yes	0.2983	0.4991	0.5765	0.8641	0.8330	0.5619
	no	0.3682	0.5928	0.6576	0.8555	0.8567	0.5545
45 to 64 years	yes	0.6018	0.5515	0.5596	0.6398	0.7271	0.5869
	no	0.6801	0.7016	0.6896	0.7407	0.8144	0.6967
Country of Birth							
Not USA	yes	0.3372	0.3844	0.4911	0.6436	0.5499	0.2973
	no	0.2314	0.2654	0.3647	0.5250	0.4283	0.2065
USA	yes	0.3926	0.6020	0.6680	0.8526	0.8359	0.6082
	no	0.4277	0.6850	0.6904	0.8069	0.8338	0.5741

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-21. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Pennsylvania sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.2993	0.5701	0.6413	0.7014	0.7467	0.4299
	no	0.3916	0.5928	0.6209	0.7428	0.7834	0.4912
Gender							
Male	yes	0.4637	0.4317	0.6876	0.6802	0.6609	0.3946
	no	0.5102	0.4556	0.7162	0.7379	0.7599	0.4231
Female	yes	0.4881	0.8425	0.7180	0.7085	0.6806	0.5616
	no	0.5130	0.8141	0.7010	0.7166	0.6174	0.5087
Race/Ethnicity							
Black	yes	0.4498	0.4837	0.4236	0.4009	0.6903	0.2434
	no	0.5721	0.5840	0.5292	0.5005	0.7693	0.3554
Hispanic	yes	0.0695	0.2131	0.1176	0.1901	0.0287	0.0733
	no	0.7943	0.9335	0.8754	0.9262	0.6088	0.8046
Other	yes	0.5618	0.5929	0.6340	0.6738	0.7626	0.6508
	no	0.6024	0.5565	0.5170	0.6450	0.7079	0.5631
Education Level							
No HS degree	yes	0.3013	0.5705	0.3927	0.4633	0.4218	0.3253
	no	0.4518	0.6120	0.5467	0.6499	0.5467	0.4989
HS degree	yes	0.4797	0.5708	0.5306	0.5300	0.8713	0.4303
	no	0.5813	0.6607	0.4997	0.5488	0.8109	0.5490
Some College	yes	0.4939	0.4402	0.7105	0.8400	0.7081	0.5718
	no	0.5410	0.4774	0.7027	0.8405	0.7197	0.5469
College graduate	yes	0.3710	0.4802	0.5338	0.3892	0.6238	0.7136
	no	0.3829	0.4775	0.5240	0.4491	0.6312	0.7026
Age							
16 to 24 years	yes	0.5394	0.6558	0.5223	0.4420	0.6081	0.5650
	no	0.4949	0.5345	0.4357	0.3664	0.5594	0.5528
25 to 44 years	yes	0.3370	0.6467	0.6743	0.8099	0.8003	0.4202
	no	0.4449	0.6782	0.6825	0.8187	0.8176	0.4414
45 to 64 years	yes	0.5317	0.4063	0.5471	0.6982	0.5934	0.5460
	no	0.6354	0.5229	0.6185	0.8032	0.6513	0.6579
Country of Birth							
Not USA	yes	0.2846	0.2855	0.3997	0.7171	0.5197	0.2930
	no	0.1936	0.1927	0.2808	0.6027	0.3897	0.2014
USA	yes	0.4150	0.5887	0.6967	0.7119	0.7624	0.4917
	no	0.4722	0.6080	0.6791	0.7579	0.7899	0.5420

* Compositing factors reflect imputation variance as well as sampling variance.

602

Table O-22. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Texas sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.4401	0.5381	0.3025	0.4647	0.6946	0.3590
	no	0.5373	0.5241	0.3884	0.4871	0.7312	0.4638
Gender							
Male	yes	0.4365	0.4458	0.3484	0.4892	0.6993	0.3559
	no	0.4793	0.4720	0.4018	0.5184	0.7619	0.4216
Female	yes	0.5817	0.4368	0.5443	0.6365	0.5172	0.4704
	no	0.6184	0.3544	0.6008	0.6734	0.5214	0.4667
Race/Ethnicity							
Black	yes	0.3481	0.3319	0.3773	0.2054	0.5934	0.1707
	no	0.4209	0.3765	0.4901	0.2305	0.6485	0.2503
Hispanic	yes	0.3100	0.4320	0.4107	0.7014	0.4118	0.2847
	no	0.5074	0.6081	0.6972	0.8387	0.5200	0.3974
Other	yes	0.6462	0.5723	0.4158	0.5066	0.7568	0.5179
	no	0.6655	0.5395	0.4324	0.4973	0.7536	0.5445
Education Level							
No HS degree	yes	0.5186	0.6386	0.5824	0.4018	0.4449	0.3181
	no	0.5620	0.7063	0.7027	0.4783	0.5139	0.3984
HS degree	yes	0.4450	0.3545	0.4565	0.5014	0.6338	0.3770
	no	0.5887	0.4258	0.5157	0.5947	0.7175	0.5421
Some college	yes	0.3858	0.4708	0.2704	0.4724	0.7016	0.4298
	no	0.4927	0.5691	0.4570	0.5373	0.7445	0.5288
College graduate	yes	0.5020	0.6807	0.6250	0.4894	0.7008	0.6200
	no	0.4574	0.6859	0.5779	0.4533	0.6929	0.5917
Age							
16 to 24 years	yes	0.3372	0.3110	0.3958	0.6546	0.3672	0.3393
	no	0.3991	0.3336	0.5032	0.7026	0.4998	0.5048
25 to 44 years	yes	0.3967	0.6324	0.3893	0.4720	0.7384	0.4007
	no	0.4470	0.5864	0.4234	0.4461	0.7482	0.4052
45 to 64 years	yes	0.4427	0.5188	0.5261	0.5361	0.4740	0.4378
	no	0.5640	0.5561	0.5410	0.5849	0.5155	0.5414
Country of Birth							
Not USA	yes	0.3307	0.4881	0.5737	0.7932	0.3848	0.2885
	no	0.4358	0.5645	0.6900	0.8386	0.5096	0.3980
USA	yes	0.4847	0.5998	0.3325	0.4344	0.6650	0.4294
	no	0.5859	0.5883	0.4007	0.4462	0.6888	0.4490

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-23. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Texas sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.3289	0.5877	0.6021	0.8470	0.7856	0.4302
	no	0.3807	0.6675	0.6518	0.8215	0.8005	0.4872
Gender							
Male	yes	0.4216	0.5078	0.5764	0.6979	0.5525	0.4404
	no	0.4063	0.5407	0.5894	0.6889	0.5900	0.4718
Female	yes	0.3462	0.6287	0.4100	0.6525	0.6864	0.4531
	no	0.4390	0.6614	0.5058	0.6701	0.6893	0.4368
Race/Ethnicity							
Black	yes	0.2817	0.4027	0.3892	0.3046	0.2146	0.1863
	no	0.3411	0.4499	0.5023	0.3397	0.2269	0.2619
Hispanic	yes	0.2895	0.4103	0.4047	0.5005	0.5993	0.2495
	no	0.3494	0.5311	0.4879	0.6379	0.6777	0.3517
Other	yes	0.5107	0.6142	0.6634	0.8118	0.8147	0.6007
	no	0.5447	0.6628	0.6829	0.7743	0.8126	0.6057
Education Level							
No HS degree	yes	0.4225	0.5119	0.5625	0.4516	0.3526	0.3598
	no	0.4542	0.5416	0.6136	0.5243	0.4166	0.4261
HS degree	yes	0.3409	0.4104	0.4964	0.5182	0.5129	0.3930
	no	0.4589	0.5304	0.5842	0.6227	0.6424	0.5541
Some college	yes	0.2857	0.6572	0.8040	0.5851	0.6697	0.4407
	no	0.3312	0.7063	0.8035	0.6291	0.6310	0.4832
College graduate	yes	0.5585	0.6557	0.4389	0.7937	0.7507	0.6562
	no	0.5260	0.6071	0.4584	0.7876	0.7642	0.6598
Age							
16 to 24 years	yes	0.4011	0.7459	0.6615	0.4617	0.4956	0.3325
	no	0.4642	0.7183	0.7383	0.5174	0.5756	0.4641
25 to 44 years	yes	0.2272	0.4079	0.4849	0.8147	0.7752	0.4699
	no	0.3017	0.5017	0.5485	0.7914	0.7916	0.4616
45 to 64 years	yes	0.5625	0.5114	0.5196	0.6018	0.6939	0.5473
	no	0.6063	0.6017	0.6032	0.6688	0.7438	0.6138
Country of Birth							
Not USA	yes	0.4075	0.4577	0.5661	0.7094	0.6229	0.3639
	no	0.4508	0.4877	0.6354	0.8045	0.7321	0.4693
USA	yes	0.2936	0.4931	0.5642	0.7881	0.7661	0.4996
	no	0.3584	0.5884	0.6034	0.7529	0.7667	0.4831

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-24. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Texas sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	0.2241	0.4726	0.5472	0.6135	0.6658	0.3376
	no	0.3454	0.5131	0.5272	0.6624	0.7115	0.4151
Gender							
Male	yes	0.3876	0.3573	0.6171	0.6089	0.5879	0.3230
	no	0.4302	0.3666	0.6153	0.6481	0.6637	0.3382
Female	yes	0.3777	0.7730	0.6185	0.6074	0.5756	0.4491
	no	0.4636	0.7680	0.6340	0.6523	0.5456	0.4512
Race/Ethnicity							
Black	yes	0.3334	0.3643	0.3101	0.2905	0.5769	0.1644
	no	0.4562	0.4302	0.3640	0.3379	0.6315	0.2653
Hispanic	yes	0.2444	0.5400	0.3662	0.5042	0.1135	0.2552
	no	0.3852	0.6590	0.5920	0.8300	0.1179	0.4480
Other	yes	0.5096	0.5414	0.5841	0.6261	0.7225	0.6017
	no	0.5785	0.5578	0.5488	0.6421	0.7186	0.5857
Education Level							
No HS degree	yes	0.2382	0.4906	0.3193	0.3850	0.3460	0.2591
	no	0.3514	0.4326	0.4258	0.5265	0.3745	0.3968
HS degree	yes	0.3371	0.4231	0.3840	0.3835	0.7888	0.2941
	no	0.5121	0.5838	0.4519	0.4797	0.7985	0.4664
Some college	yes	0.3528	0.3052	0.5782	0.7458	0.5754	0.4272
	no	0.4475	0.3923	0.6221	0.7843	0.6441	0.4500
College graduate	yes	0.3963	0.5069	0.5602	0.4148	0.6485	0.7349
	no	0.3674	0.4631	0.5031	0.4337	0.6166	0.6912
Age							
16 to 24 years	yes	0.3403	0.4563	0.3251	0.2587	0.4060	0.3640
	no	0.4795	0.5121	0.4123	0.3416	0.5453	0.5427
25 to 44 years	yes	0.2600	0.5587	0.5887	0.7465	0.7349	0.3338
	no	0.3734	0.5810	0.5848	0.7420	0.7426	0.3552
45 to 64 years	yes	0.4914	0.3680	0.5069	0.6632	0.5540	0.5058
	no	0.5559	0.4260	0.5262	0.7337	0.5554	0.5849
Country of Birth							
Not USA	yes	0.3496	0.3506	0.4737	0.7740	0.5939	0.3591
	no	0.4308	0.4279	0.5195	0.8248	0.6355	0.4511
USA	yes	0.3133	0.4794	0.5964	0.6137	0.6736	0.3836
	no	0.3972	0.5195	0.5759	0.6609	0.7061	0.4372

* Compositing factors reflect imputation variance as well as sampling variance.

Table O-25. NALS compositing analysis: Compositing factors* for average prose proficiency and literacy levels in Washington sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	**National Certainty PSU?**	**State Compositing Factor: Beta**					
		Level 1 225 or lower	**Level 2 226 to 275**	**Level 3 276 to 325**	**Level 4 326 to 375**	**Level 5 376 or higher**	**Overall Proficiency**
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.8765	0.8587	0.7978	0.8601	0.9465	0.8524
Gender							
Male	yes	–	–	–	–	–	–
	no	0.8315	0.8525	0.7959	0.8618	0.9576	0.7963
Female	yes	–	–	–	–	–	–
	no	0.9013	0.7357	0.9063	0.9349	0.8685	0.8185
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.7669	0.7446	0.7970	0.6002	0.8971	0.5728
Hispanic	yes	–	–	–	–	–	–
	no	0.7954	0.8691	0.8658	0.9539	0.8512	0.7594
Other	yes	–	–	–	–	–	–
	no	0.9027	0.8084	0.7760	0.8057	0.9292	0.8557
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.9085	0.9454	0.9442	0.8718	0.8898	0.8319
HS degree	yes	–	–	–	–	–	–
	no	0.9125	0.8327	0.8782	0.9084	0.9480	0.8973
Some college	yes	–	–	–	–	–	–
	no	0.9200	0.9078	0.8991	0.9133	0.9443	0.8952
College graduate	yes	–	–	–	–	–	–
	no	0.7330	0.9049	0.8132	0.7425	0.9070	0.8391
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.8550	0.7768	0.8968	0.9476	0.8979	0.9136
25 to 44 years	yes	–	–	–	–	–	–
	no	0.8080	0.8726	0.7965	0.8122	0.9407	0.7682
45 to 64 years	yes	–	–	–	–	–	–
	no	0.9205	0.9052	0.9052	0.9245	0.8950	0.9163
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.8897	0.9327	0.9592	0.9817	0.9193	0.8690
USA	yes	–	–	–	–	–	–
	no	0.9023	0.8860	0.8128	0.8328	0.9308	0.8228

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-26. NALS compositing analysis: Compositing factors* for average document proficiency and literacy levels in Washington sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.7587	0.9316	0.9200	0.9558	0.9617	0.8460
Gender							
Male	yes	–	–	–	–	–	–
	no	0.7568	0.8611	0.8698	0.9071	0.8940	0.8137
Female	yes	–	–	–	–	–	–
	no	0.8191	0.9152	0.8762	0.9264	0.9324	0.8022
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.6999	0.7968	0.8075	0.7162	0.6060	0.5935
Hispanic	yes	–	–	–	–	–	–
	no	0.7547	0.8508	0.8413	0.8937	0.9237	0.7216
Other	yes	–	–	–	–	–	–
	no	0.8452	0.9201	0.9168	0.9135	0.9508	0.8623
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.8531	0.9036	0.9242	0.8852	0.8450	0.8410
HS degree	yes	–	–	–	–	–	–
	no	0.8536	0.8956	0.9169	0.9117	0.9388	0.8915
Some college	yes	–	–	–	–	–	–
	no	0.7972	0.9422	0.9534	0.9101	0.8926	0.8570
College graduate	yes	–	–	–	–	–	–
	no	0.7958	0.8268	0.7990	0.9378	0.9482	0.8896
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.8771	0.9420	0.9584	0.8893	0.9181	0.9002
25 to 44 years	yes	–	–	–	–	–	–
	no	0.6915	0.8603	0.8928	0.9455	0.9557	0.7925
45 to 64 years	yes	–	–	–	–	–	–
	no	0.9178	0.9385	0.9282	0.9402	0.9636	0.9286
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.8804	0.9013	0.9453	0.9786	0.9670	0.8959
USA	yes	–	–	–	–	–	–
	no	0.7438	0.9104	0.8987	0.9328	0.9504	0.8276

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Table O-27. NALS compositing analysis: Compositing factors* for average quantitative proficiency and literacy levels in Washington sample, by Total Population, Gender, Race/Ethnicity, Education Level, Age, and Country of Birth

Demographic Subpopulations	National Certainty PSU?	State Compositing Factor: Beta					
		Level 1 225 or lower	Level 2 226 to 275	Level 3 276 to 325	Level 4 326 to 375	Level 5 376 or higher	Overall Proficiency
Total Population							
Total	yes	–	–	–	–	–	–
	no	0.7425	0.8592	0.8644	0.9323	0.9449	0.8149
Gender							
Male	yes	–	–	–	–	–	–
	no	0.7796	0.7310	0.8931	0.9109	0.9326	0.7066
Female	yes	–	–	–	–	–	–
	no	0.8327	0.9504	0.9178	0.9287	0.8730	0.8132
Race/Ethnicity							
Black	yes	–	–	–	–	–	–
	no	0.7681	0.7744	0.7304	0.7082	0.8900	0.5771
Hispanic	yes	–	–	–	–	–	–
	no	0.7296	0.9058	0.8372	0.9071	0.5159	0.7488
Other	yes	–	–	–	–	–	–
	no	0.8901	0.8568	0.8153	0.9030	0.9165	0.8408
Education Level							
No HS degree	yes	–	–	–	–	–	–
	no	0.7792	0.8593	0.8388	0.8969	0.8380	0.8153
HS degree	yes	–	–	–	–	–	–
	no	0.8776	0.9153	0.8409	0.8732	0.9543	0.8688
Some college	yes	–	–	–	–	–	–
	no	0.8800	0.8334	0.9109	0.9603	0.9250	0.8319
College graduate	yes	–	–	–	–	–	–
	no	0.6717	0.7384	0.7608	0.7674	0.8469	0.8748
Age							
16 to 24 years	yes	–	–	–	–	–	–
	no	0.8920	0.8761	0.8511	0.8246	0.9098	0.9250
25 to 44 years	yes	–	–	–	–	–	–
	no	0.7746	0.8793	0.8726	0.9358	0.9381	0.7236
45 to 64 years	yes	–	–	–	–	–	–
	no	0.9033	0.8645	0.9004	0.9604	0.9140	0.9100
Country of Birth							
Not USA	yes	–	–	–	–	–	–
	no	0.8800	0.8828	0.9115	0.9789	0.9407	0.8886
USA	yes	–	–	–	–	–	–
	no	0.7834	0.8582	0.8873	0.9340	0.9410	0.8349

* Compositing factors reflect imputation variance as well as sampling variance.

– Not available.

Appendix P

ELECTRONIC CODE BOOK FOR WINDOWS USER'S MANUAL

Appendix P: Electronic Code Book for Windows User's Manual
NATIONAL ADULT LITERACY SURVEY

When the Electronic Code Book for Windows (ECBW) is opened, the user will see the main window containing variable names and label names for the 1992 National Adult Literacy Survey. At the top of the table is a menu, and several buttons, which allows users to move throughout the ECBW and obtain the information needed. A description of all of the menu options is provided below. Shortcuts and toolbar buttons are underlined and described within the menu text.

File Menu

The File menu options provide the user with the ability to move and export information about the variables selected from the main window.

- **Output.** The Output option of the File menu allows the user to export the codes that have been selected from the table. The user may export this information to one of the following types of files:

    1. *SAS-PC Code*—Allows the user to create a file that tells a SAS program how to read the data;

    2. *SPSS*—Allows the user to create a file that tells a SPSS program how to read in the data. The data in this file is compressed;

    3. *Code Book Text*—Allows the user to create a file that contains information about the variables selected (frequency, labels, etc.);

    4. *Tag File*—Allows the user to create a file that contains the variables selected so that items do not have to be re-selected each time the user enters the program.

    5. *AccessDB*—Allows the user to create an Access database file that contains the variables selected.

- **View Output.** The View Output option allows the user to look at the output file created

- **Import Tag Files.** The Import Tag Files option allows the user to recall a previously created tag file.

- **Set Up.** The Set Up option tells the user what directory the files are in and where the files are located.

- **Exit.** Used to exit the ECBW.

Move Menu

The Move menu allows the user to move between variable names and labels within the ECBW.

The menu options are:

- Top of List
- Prev Section (previous section)
- Next Section
- Bottom of List
- Prev File (previous file)
- Next File

The six black arrow buttons located on the toolbar can also be used to move within ECBW without accessing the Move menu.

Move Button Options:

Arrow Button	**Function**
Left/Right arrows with double lines:	Move to top of list/bottom of list
Left/Right arrows with single line:	Move to previous file/next file
Left/Right single arrows:	Move to previous section/next section

Tag Menu

The Tag menu provides options for selecting and deselecting items.

- **Tag/Untag Items.** The Tag Items option in the Tag menu allows the user to select variables within the ECBW once they are highlighted with the mouse. The Untag Items option allows the user to deselect individual items. Users may also select a variable by simply clicking in the box located next to each variable. To deselect the user must click the box a second time.

- **Clear All Tags.** The Clear All Tags option erases all the checked boxes in order to make a new selection of items.

- **Previous Tag/Next Tag.** The Previous Tag and Next Tag options allow the user to move back and forth between the selected variables. Users may also move between tags by clicking on the blue and red arrow buttons located on the toolbar.

View Menu

Once a variable is selected, the user may choose to obtain a more detailed description of the variable. The View menu provides this information.

- **Description.** When chosen, this option produces a Description/Frequency Window. Users may also reach the Description/Frequency Window by double clicking on the selected variable.

- **Description/Frequency Window.** This window has two options.

  1. *View Description.* The header over the window provides the user with information about the survey and where the information came from. The text inside the window provides a description of the variable (parameters, how it was derived, etc.).

  2. *View Frequency.* Provides the user with the code, frequency, and percent category label of the variable selected.

     To exit the window, the user must click on the *X* button in the top right corner of the Description/Frequency Window.

- **Tagged Items.** This option in the View menu allows the user to create a list of the items that have been tagged/checked from the main window. The user may also create a list of tagged items by clicking on the toolbar button that looks like a sheet of paper.

Search Menu

When the Search menu is selected, a search window will appear. This window allows the user to search the Code Book by variable, label, or description in a forward or backward direction. To exit this window, the user must click on the X button in the top right corner of the search window. This window can also be accessed by clicking on the magnifying glass button located on the toolbar.

Help Menu

The Help menu (located in the far right hand corner of the main window) provides the user with information pertaining to the ECBW.

- **Contents.** The user may search for information on a certain topic by selecting the Contents option. Once selected, a new window will appear which allows the user to select a topic of interest.

Help topics include:

- System overview
- The ECB Main Window
- System Requirements
- How to Tag/Untag a Variable
- How to View Descriptions/Frequencies
- How to Create SAS/SPSS Program Code, Codebook Text, Tag and Access Database Files
- How to Navigate Through Variables
- How to View Tagged Variables
- How to Import Tag Files
- How to Search for Text
- How to Change the System Setup

- **Search**—If the user is unsure of which topic to select, or would prefer to search for specific words and phrases in the Help feature, the user must click on the Search button at the top of the Help screen, and follow the given instructions. The user may choose to perform either an Index search, or a Find search by clicking on the tabs located at the top left of the Search Window.

- **Back**—The Back button brings the user back to main topics page, once a topic has been selected and viewed.

- **Print**—Allows the user to print out selections of the Help Manual for hard copy.

- **Other Menu Options: File, Edit, Bookmark, Options, Help.**
  These features allow the user to further manipulate the information in the Help Manual. Certain pages may be copied, saved to another file, bookmarked or annotated, if needed.

To exit the Help window, the user must click on the X button located on the top right corner of the Help window. The user is able to keep this window open while working within the ECBW by shrinking it on the desktop. To minimize the window, the user must click the dashed line button, also located in the top right corner of the Help window (to the left of the X). To enlarge the window, the user must click on the view window button, which is located to the left of the close window (X) button when minimized at the bottom of the screen.

Appendix Q

STANDARD ERRORS FOR TABLE 8.2

Appendix Q: Standard Errors for Table 8.2
NATIONAL ADULT LITERACY SURVEY

Corrected with Taylor Series

		Sufficient cognitive data present				Insufficient cognitive data present			
		Reasons Related to Literacy				Reasons unrelated to literacy			
		Total	Non-English Language	Mental Disability	Reading or writing difficulty	Total	Refused	Physical Disability	Other Unknown No answer
Total	0.005	0.003	0.002	0.001	0.002	0.004	0.002	0.002	0.001
Age									
16 to 64	0.004	0.002	0.002	0.001	0.001	0.003	0.002	0.001	0.001
65 and older	0.015	0.008	0.005	0.004	0.006	0.014	0.008	0.008	0.005
Race/ethnicity									
Hispanic	0.015	0.015	0.015	0.001	0.004	0.005	0.003	0.002	0.003
Asian	0.032	0.028	0.027	0.007	0.004	0.016	0.007	0.012	0.008
Black	0.008	0.006	0.003	0.003	0.004	0.006	0.004	0.004	0.003
White	0.005	0.002	0.000	0.001	0.002	0.004	0.003	0.002	0.002
Other or missing	0.037	0.035	0.029	0.000	0.011	0.016	0.006	0.011	0.012
Language spoken while growing up									
English	0.004	0.002	0.000	0.001	0.002	0.004	0.003	0.002	0.001
Language other than English	0.017	0.015	0.014	0.003	0.006	0.006	0.004	0.004	0.004
Missing data	0.052	0.008	0.000	0.003	0.007	0.051	0.000	0.000	0.051
English spoken while growing up									
Hispanic	0.009	0.006	0.004	0.001	0.004	0.008	0.005	0.002	0.004
Asian	0.042	0.000	0.000	0.000	0.000	0.042	0.016	0.037	0.016
Black	0.008	0.005	0.000	0.003	0.004	0.006	0.004	0.004	0.003
White	0.005	0.002	0.000	0.001	0.002	0.004	0.003	0.002	0.002
English not spoken while growing up									
Hispanic	0.019	0.019	0.020	0.001	0.006	0.007	0.004	0.003	0.005
Asian	0.039	0.037	0.036	0.010	0.006	0.011	0.007	0.005	0.009
Black	0.082	0.085	0.087	0.019	0.000	0.019	0.014	0.003	0.014
White	0.031	0.027	0.017	0.011	0.018	0.016	0.011	0.011	0.010
Education									
0 to 8 years	0.019	0.014	0.014	0.005	0.010	0.013	0.007	0.009	0.007
9 to 12 years	0.011	0.006	0.005	0.003	0.004	0.009	0.005	0.005	0.004
HS grad or GED	0.006	0.003	0.002	0.001	0.002	0.005	0.004	0.002	0.002
Some college	0.004	0.002	0.001	0.001	0.001	0.003	0.003	0.002	0.001
College degree	0.005	0.001	0.001	0.001	0.000	0.004	0.003	0.001	0.002
Other or Missing	0.014	0.012	0.012	0.003	0.002	0.005	0.004	0.003	0.003

Appendix R

STANDARD ERROR TABLES FOR CHAPTER 10

Appendix R: Standard Error Tables for Table 10
NATIONAL ADULT LITERACY SURVEY

Chapter 10 **(se's of proportions based on experimental design)**

Table 10-2. Standard error of response rate

	$0	$20	$35	overall
Background questionnaire				
Response rate	0.786	0.827	0.844	0.836
N	730	740	818	1558
SE	0.015	0.014	0.013	0.009
Exercise booklet				
Response rate	0.928	0.979	0.98	0.980
N	574	612	690	1302
SE	0.011	0.006	0.005	0.004

Background questionnaire (t-tests)

	t of row difference
$0-$20	2
$20-$35	0.9

Exercise questionnaire (t-tests)

	t of row difference
$0-$20/$35	4.5
$20-$35	0.1

Table 10-3. Standard errors (SE)
(based on assumed equal distribution of cases across all incentive situations within a row)

	Background Questionnaire						Exercise Booklet					
	$0	**$20**	**$35**	**overall**	**t $0-$20**	**t $0-$35**	**$0**	**$20**	**$35**	**overall**	**t $0-$20**	**t $0-$35**
Total age (SE)	0.015	0.014	0.013	0.009			0.011	0.006	0.005	0.004		
N	331	331	331	331	0.00		287	287	287	287		
Response rate	0.8	0.884	0.902	0.867			1	1	1	1		
16-24 (SE)	0.038	0.030	0.028	0.019	1.72	2.15	NA	NA	NA	NA		
N	959	959	959	959			807	807	807	807		
Response rate	0.784	0.873	0.86	0.842			0.944	0.983	0.996	0.976		
25-44 (SE)	0.023	0.019	0.019	0.012	3.01	2.52	0.014	0.008	0.004	0.005	2.43	3.58
N	534	534	534	534			423	423	423	423		
Response rate	0.749	0.79	0.833	0.792			0.896	0.958	0.962	0.94		
45-64 (SE)	0.032	0.031	0.028	0.018	0.92	1.96	0.026	0.017	0.016	0.012	2.02	2.18
N	253	253	253	253			206	206	206	206		
Response rate	0.837	0.758	0.831	0.814			0.934	0.978	0.953	0.952		
65-74 (SE)	0.040	0.047	0.041	0.024	-1.28	-0.10	0.030	0.018	0.026	0.015	1.26	0.48
N	160	160	160	160			130	130	130	130		
Response rate	0.86	0.736	0.842	0.813			0.743	0.97	0.919	0.876		
75+ (SE)	0.048	0.060	0.050	0.031	-1.61	-0.26	0.066	0.026	0.041	0.029	3.19	2.25
N	51	51	51	51			23	23	23	23		
Response rate	0.545	0.357	0.462	0.451			1	0.33	1	0.875		
Missing	0.121	0.116	0.121	0.070	-1.12	-0.49	NA	0.170	NA	0.069		

Table 10-3. Standard errors — Continued
(based on assumed equal distribution of cases across all incentive situations within a row)

	Background Questionnaire						Exercise Booklet					
	$0	$20	$35	overall	t $0-$20	t $0-$35	$0	$20	$35	overall	t $0-$20	t $0-$35
Race/ethnicity												
N	1769	1769	1769	1769			1448	1448	1448	1448		
Response rate	0.787	0.811	0.853	0.819			0.934	0.975	0.979	0.964		
White (SE)	0.017	0.016	0.015	0.009	1.03	2.96	0.011	0.007	0.007	0.005	3.07	3.45
N	226	226	226	226			187	187	187	187		
Response rate	0.754	0.883	0.833	0.827			0.864	0.985	0.971	0.949		
Black (SE)	0.050	0.037	0.043	0.025	2.08	1.20	0.043	0.015	0.021	0.016	2.63	2.21
N	161	161	161	161			143	143	143	143		
Response rate	0.857	0.914	0.894	0.888			0.917	1	1	0.969		
Hispanic (SE)	0.048	0.038	0.042	0.025	0.93	0.58	0.040	NA	NA	0.014		
N	69	69	69	69			62	62	62	62		
Response rate	0.8	0.958	0.84	0.87			0.909	1	1	0.98		
Other (SE)	0.083	0.042	0.076	0.040	1.69	0.35	0.063	NA	NA	0.018		
N	63	63	63	63			36	36	36	36		
Response rate	0.667	0.571	0.571	0.603			0.917	0.667	1	0.9		
Missing	0.103	0.108	0.108	0.062	-0.64	-0.64	0.080	0.136	NA	0.050	-1.59	

Table 10-3. Standard errors — Continued
(based on assumed equal distribution of cases across all incentive situations within a row)

	Background Questionnaire						Exercise Booklet					
	$0	$20	$35	overall	t $0-$20	t $0-$35	$0	$20	$35	overall	t $0-$20	t $0-$35
Education												
N							68	68	68	68		
							1	0.955	1	0.985		
Still in HS (SE)							NA	0.044	NA	0.015		
N							358	358	358	358		
Response rate							0.721	0.78	0.789	0.767		
No HS diploma (SE)							0.041	0.038	0.037	0.022	1.06	1.23
N							486	486	486	486		
Response rate							0.867	0.949	0.966	0.894		
HS diploma (SE)							0.027	0.017	0.014	0.014	2.58	3.27
N							553	553	553	553		
Response rate							0.923	0.952	0.973	0.951		
Some college (SE)							0.020	0.016	0.012	0.009		
N							345	345	345	345		
Response rate							0.952	0.962	0.982	0.965		
College degree (SE)							0.020	0.018	0.012	0.010		
N							66	66	66	66		
Response rate							0.467	1	0.65	0.645		
Missing							0.106	NA	0.102	0.059		

Note: NA indicates no variation in (100% response rate) in cell

Table 10-3. Standard errors for exercise booklet (no high school vs. high school)

		No High Sch	High School	t of row
SE	$0	0.041	0.027	2.980
	$20	0.038	0.017	4.055
	$35	0.037	0.014	4.428

Table 10-3. Standard errors for exercise booklet (some college vs. college)

		Some College	College	t of row
SE	$0	0.020	0.020	1.036
	$20	0.016	0.018	0.420
	$35	0.012	0.012	0.523

	t $20-$35
Some College	0.420
College	1.063

Table 10-3. Standard errors (minority population equals black and Hispanic population combined)

Background questionnaire

		Response Rate HisBl	SE HisBlack	Response Rate White	SE White	t of row		t$20-$35
SE	$0	0.797	0.035	0.787	0.017	0.257	**HisBlack**	0.929
	$20	0.896	0.027	0.811	0.016	2.708		
	$35	0.858	0.031	0.853	0.015	0.158		

Exercise booklet

		Response Rate HisBl	SE HisBlack	Response Rate White	SE White	t of row		t $20-$35	t $0-$20
SE	$0	0.887	0.03	0.934	0.011	1.46	**HisBlack**	0.53	3.33
	$20	0.992	0.009	0.975	0.007	1.46			
	$35	0.984	0.012	0.979	0.007	0.332			

Note: HisBl represents the combination of the Black and Hispanic poulations

Table 10-3. Standard errors (ages 16-64 combined and 65-75+ combined)

Background questionnaire

		Response Rate 16-64	SE 16-64	Response Rate 65-75+	SE 65-75+	t of row
SE	$0	0.777	0.017	0.844	0.031	1.91
	$20	0.851	0.0145	0.758	0.037	2.6
	$35	0.86	0.014	0.833	0.032	0.762

Exercise booklet

		Response Rate 16-64	SE 16-64	Response Rate 65-75+	SE 65-75+	t of row
SE	$0	0.941	0.011	0.86	0.033	2.34
	$20	0.979	0.006	0.975	0.015	0.229
	$35	0.987	0.005	0.94	0.022	2.031

T-tests for background questionnaire compared to exercise booklet

	Response Rate 16-64	Response Rate 65-75+
$0	3.011	0.361
$20	8.085	5.707
$35	8.471	2.742

Table 10-4. Standard errors for initial cooperation rate, by incentive level: 1991 Field test

Incentive	**$0**	**$20**	**$35**			
N	566	603	684			
Response rate	0.126	0.066	0.051	**t of $0-$20**	**t of $0-$35**	**t of $25-$35**
Incomplete (SE)	0.014	0.010	0.008	-3.48	-4.60	-1.14

Table 10-5. Standard errors for initial cooperation rate, by incentive level: Older adults 1991 field test

Incentive	**$0**	**$20**	**$35**			
N	135	88	116			
Response rate	0.24	0.09	1	**t of $0-$20**	**t of $0-$35**	**t of $25-$35**
Incomplete (SE)	0.037	0.031	0.028	-3.14	-3.04	0.24

Table 10-7. Standard errors for prose scale

	Total Std Error	$0 incentive Std Error	$20 incentive Std Error	$35 incentive Std Error	t $0-$20	t $0-$35	t $20-$35
Total	0.244	0.454	0.417	0.396	3.407	2.673	0.869
Sex							
Male	0.366	0.647	0.626	0.617			
Female	0.330	0.640	0.558	0.520			
Race/ethnicity							
White	0.270	0.497	0.481	0.432			
Black	0.704	1.553	0.968	1.190			
Hispanic	0.853	1.418	1.401	1.483			
Education							
Still in H.S.	1.249	2.840	2.226	1.578			
< HS	1.031	1.856	1.734	1.553			
Some HS	0.693	1.342	1.042	1.278			
GED/HSEQ	1.266	1.869	1.897	2.291			
HS diploma	0.447	0.990	0.685	0.697			
Some college	0.430	0.769	0.760	0.706			
College degree	0.411	0.661	0.702	0.733			
Age							
16-20	0.785	1.643	1.241	1.225			
21-25	0.792	1.358	1.729	1.108			
26-31	0.620	1.212	1.071	0.911			
32-45	0.420	0.843	0.661	0.695			
46-64	0.519	0.936	0.884	0.850	**t $0-$20**	**t $0-$35**	**t $20-$35**
65+	0.626	1.005	1.096	1.059	3.43	3.014	0.459
Income							
<$5,000	1.088	2.125	1.450	2.023			
$5-9,999	1.121	2.012	1.928	1.683			
$10-14,999	0.831	1.793	1.489	1.172			
$15-19,999	0.902	1.253	1.677	1.472			
$20-29,999	0.657	1.174	1.120	1.110			
$30-39,999	0.637	1.428	1.017	0.966			
$40-49,999	0.602	1.046	1.059	0.975			
$50,000+	0.422	0.742	0.776	0.673			
Refused	1.708	2.935	2.432	3.033			
Don't know	1.313	2.132	2.013	2.400			

Table 10-8. Standard errors for document scale

	Total Std Error	$0 incentive Std Error	$20 incentive Std Error	$35 incentive Std Error	t $0-$20	t $0-$35	t $20-$35
Total	0.244	0.440	0.417	0.396	2.97	3.55	0.522
Sex							
Male	0.370	0.640	0.632	0.629			
Female	0.324	0.615	0.558	0.514			
Race/ethnicity							
White	0.265	0.482	0.476	0.424			
Black	0.696	1.361	1.084	1.131			
Hispanic	0.970	1.479	1.446	1.967			
Education							
Still in H.S.	1.273	2.482	2.008	2.040			
< HS	1.166	2.482	1.990	1.704			
Some HS	0.641	1.254	0.960	1.175			
GED/HSEQ	1.157	1.447	1.578	2.357			
HS diploma	0.427	0.812	0.742	0.665			
Some college	0.430	0.786	0.813	0.660			
College degree	0.426	0.744	0.746	0.717			
Age							
16-20	0.777	1.477	1.285	1.238			
21-25	0.671	1.157	1.073	1.157			
26-31	0.640	1.285	1.071	0.989			
32-45	0.386	0.731	0.640	0.644			
46-64	0.541	1.015	0.971	0.858	**t $0-$20**	**t $0-$35**	**t $20-$35**
65+	0.596	0.936	1.152	0.961	2.133	3.578	0.668
Income							
<$5,000	1.332	2.700	1.727	2.315			
$5-9,999	1.020	1.856	1.739	1.556			
$10-14,999	0.806	1.706	1.180	1.272			
$15-19,999	0.822	1.236	1.473	1.371			
$20-29,999	0.657	1.152	1.174	1.076			
$30-39,999	0.611	1.248	0.962	0.966			
$40-49,999	0.637	1.108	1.132	1.023			
$50,000+	0.413	0.758	0.719	0.681			
Refused	1.802	2.838	3.154	3.267			
Don't know	1.275	1.897	3.056	1.860			

Table 10-9. Standard errors for quantitative scale

	Total Std Error	$0 incentive Std Error	$20 incentive Std Error	$35 incentive Std Error	t $0-$20	t $0-$35	t $20-$35
Total	0.240	0.418	0.443	0.381	3.284	3.713	0.171
Sex							
Male	0.363	0.653	0.638	0.593			
Female	0.314	0.535	0.605	0.483			
Race/ethnicity							
White	0.257	0.437	0.471	0.419			
Black	0.665	1.377	1.175	0.966			
Hispanic	0.871	1.479	1.506	1.317			
Education							
Still in H.S.	1.188	2.102	2.357	1.751			
< HS	1.052	1.677	1.972	1.598			
Some HS	0.626	1.065	0.995	1.149			
GED/HSEQ	0.912	1.508	1.599	1.375			
HS diploma	0.447	0.871	0.792	0.681			
Some college	0.416	0.718	0.795	0.652			
College degree	0.411	0.686	0.772	0.678			
Age							
16-20	0.745	1.342	1.559	1.038			
21-25	0.792	1.481	1.759	0.972			
26-31	0.633	1.184	1.038	1.022			
32-45	0.412	0.723	0.725	0.665			
46-64	0.519	0.887	1.020	0.800	**t $0-$20**	**t $0-$35**	**t $20-$35**
65+	0.602	0.946	1.073	1.030	2.517	3.718	1.076
Income							
<$5,000	1.434	3.075	2.324	2.210			
$5-9,999	0.929	1.789	1.569	1.414			
$10-14,999	0.873	1.619	1.519	1.335			
$15-19,999	0.963	1.602	1.852	1.386			
$20-29,999	0.599	1.006	1.054	1.053			
$30-39,999	0.598	1.345	0.962	0.888			
$40-49,999	0.560	0.923	1.035	0.927			
$50,000+	0.413	0.642	0.825	0.650			
Refused	1.755	2.644	2.352	4.633			
Don't know	1.088	1.556	2.110	1.520			

Table 10-13. Standard errors

Screener			
Incentive	**$0**	**$20**	**t of row difference**
Completed screener	0.009	0.002	4.36
Refusal	0.007	0.002	-2.98
Language problem	0.002	0.000	-1.90
Not at home	0.004	0.001	-1.29
Other nonresponse	0.004	0.001	-2.18

Table 10-14. Standard errors

Background questionnaire

Incentive	**$0**	**$20**	**t of row difference**
Completed BQ	0.012	0.003	6.88
Refusal	0.010	0.002	-5.32
Language problem	0.002	0.001	1.07
Physical/mental disability	0.005	0.001	-2.07
Not at home	0.006	0.001	-2.72
Other nonresponse	0.004	0.001	-1.78
Exercise booklet			
Completed exercises	0.014	0.003	7.48
Partially complete	0.009	0.002	-1.25
Language barrier	0.003	0.001	3.39
Physical/mental disability	0.005	0.001	-1.53
Reading/writing barrier	0.004	0.001	0.86
Breakoff, unwilling	0.004	0.001	-1.65
Other nonresponse	0.005	0.001	-2.30
Excerisie not started	0.012	0.002	-7.85
Language problem	0.007	0.001	-2.77
Physical/mental disability	0.005	0.001	-2.54
Reading/writing barrier	0.005	0.001	-3.02
Refusal	0.008	0.001	-5.71
Other nonresponse	0.002	0.000	-0.95

Table 10-15. Standard errors

Background questionnaire response rates

Incentive	**$0**	**$20**	**t of row difference**
All	0.015	0.003	5.90
Age			
16-24	0.036	0.007	5.16
25-44	0.019	0.004	4.40
45-64	0.024	0.006	2.67
65+	0.029	0.011	1.91
Sex			
Male	0.021	0.005	4.22
Female	0.016	0.004	5.41
Race/ethnicity			
Hispanic	0.031	0.009	0.78
Black	0.027	0.007	4.79
White and other	0.016	0.004	5.18

Table 10-16. Standard errors

Exercise booklet			
Incentive	**$0**	**$20**	**t of row difference**
All	0.011	0.002	7.29
Age			
16-24	0.024	0.003	2.24
24-44	0.016	0.002	5.00
45-64	0.021	0.004	3.34
65+	0.031	0.009	2.85
Sex			
Male	0.018	0.003	4.59
Female	0.014	0.002	5.59
Race/ethnicity			
Hispanic	0.023	0.005	1.35
Black	0.029	0.005	5.13
White and other	0.013	0.002	5.20
Education			
Some or no high school	0.024	0.005	4.49
HS diploma or GED	0.024	0.004	4.98
Some postsecondary	0.018	0.003	2.81
Some postsecondary	0.019	0.004	1.41

United States
Department of Education
ED Pubs
8242–B Sandy Court
Jessup, MD 20794–1398

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17

Standard Mail (B)